



Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCTD/DP/Gapdar  
CDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (ILML)  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nsl.nsl.nih.gov](mailto:nisc_mgc@nsl.nsl.nih.gov)

Shenchenko, Y., Wehnerby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantirpop, S., Thomas, P., Tjongson, E.B., Touchman, J.W., Tsurgon, C., Vogt, D.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILML at: <http://image.llnl.gov>  
Series: IRAL Plate: 33 Row: C Column: 17  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES  
source

Location/Qualifiers  
1..1964  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:23280 IMAGE:4637504"  
/tissue\_type="lung, large cell carcinoma"  
/clone\_id="N1H\_MGC\_18"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
277..1260  
/codon\_start=1  
/product="Unknown (protein for MGC:23280)"  
/protein\_id="AAH15582.1"  
/db\_xref="gi:15990434"

## CDS

BASE COUNT 397 a 545 c 626 g 396 t  
ORIGIN  
Query Match 84.3% Score 1635; DB 9; Length 1964;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 185; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

254 GCGCCAAACAGCGGATGAGAAAGTACGCGCTGAGCTGCGCGCGGAGAGCGCGC 313  
253 GCGCCAAACAGCGGATGAGAAAGTACGCGCTGAGCTGCGCGCGGAGAGCGCGC 312  
314 GTGTGTGCTGCTGCGGAGGAGGAGGCGGCGGCGGCTGCTGAGCTGCGGAG 373  
313 GTGTGTGCTGCTGCGGAGGAGGAGGCGGCGGCGGCTGCTGAGCTGCGGAG 372  
374 GAGAGTGGGAACATAGAGTCTTCAATGAGCTTGAAGCTTGGCACTGCGCTCGGTG 433  
373 GAGAGTGGGAACATAGAGTCTTCAATGAGCTTGAAGCTTGGCACTGCGCTCGGTG 432  
434 GCGGCTTGGCACTGCGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 493  
433 GCGGCTTGGCACTGCGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 492  
494 GCGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 553

493 GCGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 552  
554 CATATGCGTCCCTTCTGCTGACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613  
553 CATATGCGTCCCTTCTGCTGACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612  
614 AGCGCGT 673  
613 AGCGCGT 672  
674 CTGACCG 733  
673 CTGACCG 732  
734 GCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 793  
733 GCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 792  
794 TATGACG 853  
793 TATGACG 852  
854 CTGCG 913  
853 CTGCG 912  
914 CAGACACCG 973  
913 CAGACACCG 972  
974 GCGCACTGCG 1033  
973 GCGCACTGCG 1032  
1034 CTATGCG 1093  
1033 CTATGCG 1092  
1094 GAGACCG 1153  
1093 GAGACCG 1152  
1154 GAGCG 1213  
1153 GAGCG 1212  
1214 ACGGACCG 1273  
1213 ACGGACCG 1272  
1274 GATGCTGCG 1333  
1273 GATGCTGCG 1332  
1334 TGAACCTGACCG 1393  
1333 TGAACCTGACCG 1392  
1394 GCG 1453  
1393 GCG 1452  
1454 GACAGTACCG 1513  
1453 GACAGTACCG 1512  
1514 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1573  
1513 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572  
1574 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1633

Db	1573	GAGTTTGAAATGTGATGCACTGGCCCAACATTTGAGATTTGTGAACTGATTCCTTTGCACAC	1632
Qy	1634	GTCTAAGCTAGAGTATGTTAAATTAACCCCACTGTTAATGAAGCGGAATTAAGCTCCGAGCTA	1693
Db	1633	GTCTAAGCTAGAGTATGTTAAATTAACCCCACTGTTAATGAAGCGGAATTAAGCTCCGAGCTA	1692
Qy	1694	AGGAGCTCGCCTAAGGCTGTCAACAGGAATGAGAGAGGCGCTGGGATCTGAAACCCAGAGG	1755
Db	1693	AGGAGCTCGCCTAAGGCTGTCAACAGGAATGAGAGAGGCGCTGGGATCTGAAACCCAGAGG	1755
Qy	1754	TCTGAGGCGAAGGCGCGCATGCGCGTAAATGGGTCTGAGAAAGTGAAGTCAAGGCGAGGCGAG	1813
Db	1753	TCTGAGGCGAAGGCGCGCATGCGCGTAAATGGGTCTGAGAAAGTGAAGTCAAGGCGAGGCGAG	1812
Qy	1814	CTGGTATCGAAGTGCCTCCATGGAATGAAGGAGACCGCTTCGGGCGGATGCAAGGCTGGG	1873
Db	1813	CTGGTATCGAAGTGCCTCCATGGAATGAAGGAGACCGCTTCGGGCGGATGCAAGGCTGGG	1872
Qy	1874	GTCAATCTGTATCTGAAGGCGCTTCGGGAATTAAGCGGCTTGACCGCCAAAAA	1933
Db	1873	GTCAATCTGTATCTGAAGGCGCTTCGGGAATTAAGCGGCTTGACCGCCAAAAA	1932
Qy	1934	AAAAA 1939	
Db	1933	AAAAA 1938	
RESULT 2			
AX195174		1779 bp	DNA
LOCUS	AX195174		Linear
DEFINITION	Sequence 27 from Patent WO0151638.		PAT 28-AUG-2001
ACCESSION	AX195174		
VERSION	AX195174.1	GI:15385737	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE			
AUTHORS	1 Yang, J., Baughn, M.R., Burford, N., Au-Yang, J., Lu, D.A., Reddy, R.,		
	Ring, H.Z., Hillman, J.L., Yue, H., Azinza, Y., Yao, M.G., Gandhi, A.R.,		
	Nguyen, D.B., Tang, Y.T., Lai, P. and Bandman, O.		
TITLE	Drug metabolizing enzymes		
JOURNAL	Patent: WO 0151638-A 27 19-JUL-2001;		
	Incyte Genomics, Inc. (US)		
FEATURES			
source	1..1779		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/note="Incyte ID No: 2706492CB1"		
BASE COUNT	385 a 475 c 523 g 396 t		
ORIGIN			
Query Match	77.2%; Score 1497; DB 6; Length 1779;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 1547; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
Qy	371	CAGAGAGATGGGAACAATGAGGTACTCTTCAATGCGCTTGAAGCTGGCCAGCTGGCGCTCG	430
Db	222	CAGAGAGATGGGAACAATGAGGTACTCTTCAATGCGCTTGAAGCTGGCCAGCTGGCGCTCG	281
Qy	431	GTGGGGGCGCTTGGCACTGCTCTTTCTGAAGCTCTGAGCAGCGTTGGAATCTCTATCCAC	490
Db	282	GTGGGGGCGCTTGGCACTGCTCTTTCTGAAGCTCTGAGCAGCGTTGGAATCTCTATCCAC	341
Qy	491	AATGCGGATATCAGTTCTCTGTGGCGGACCCGTCAGGCGTTTAACTCTGCTGGGGTG	550
Db	342	AATGCGGATATCAGTTCTCTGTGGCGGACCCGTCAGGCGTTTAACTCTGCTGGGGTG	401
Qy	551	AACCATATCGGCGCTTCTGTGTGACATCTGCGCGCGCTTGTGCTTAAGCAATGATGCC	610
Db	402	AACCATATCGGCGCTTCTGTGTGACATCTGCGCGCGCTTGTGCTTAAGCAATGATGCC	461

[illegible]

QY 1691 CTAGGAGATCGCTGAGGCTCTCACTGATGAGAGAGAGGCTTGGGATCTGAAACCA 1750  
 1542 CTAGGAGATCGCTGAGGCTCTCACTGATGAGAGAGAGGCTTGGGATCTGAAACCA 1601  
 QY 1751 GGGGCTGAGGCGAGGCGGAGCTGCGGAGTGGGCTGAGAGAGAGGCTTGGGATCTGAAACCA 1610  
 Db 1602 GGGGCTGAGGCGAGGCGGAGCTGCGGAGTGGGCTGAGAGAGAGGCTTGGGATCTGAAACCA 1661  
 QY 1811 CAGCTGATGAGAGGCTTGGGATCTGAAACCA 1750  
 Db 1662 CAGCTGATGAGAGGCTTGGGATCTGAAACCA 1721  
 QY 1871 GGGGCTGATGAGAGGCTTGGGATCTGAAACCA 1918  
 Db 1722 GGGGCTGATGAGAGGCTTGGGATCTGAAACCA 1769

RESULT 3  
 AC024267/c 175020 bp DNA linear PRI 07-JAN-2003  
 LOCUS Homo sapiens chromosome 17, clone RP11-20824, complete sequence.  
 AC024267  
 AC024267.18 GI:27531664  
 HTG.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 175020)  
 Birren, B., Nuebaum, C. and Lander, E.  
 Homo sapiens chromosome 17, clone RP11-20824  
 2 (bases 1 to 175020)  
 Birren, B., Linton, L., Nuebaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barnes, N., Bede, F., Boguslavsky, L.,  
 Bouhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,  
 Choepel, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
 DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Donah, M., Doyle, M.,  
 Fenebor, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D.,  
 Galagan, J., Gardyna, S., Gande, S., Goyette, M., Graham, L.,  
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
 Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
 Klein, J., Landers, T., Largocque, K., Leheczy, J., Levine, R.,  
 Liew, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M.,  
 McEwan, P., McGurk, A., McKernan, K., McPeters, R., Meldrum, J.,  
 Menes, L., Mihova, T., Miranda, C., Mieng, V., Morrow, J., Naylor, J.,  
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
 Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, S.,  
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A.,  
 Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,  
 Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and  
 Zody, M.

TITLE Direct Submission  
 JOURNAL Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 175020)  
 Birren, B., Nuebaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
 Barnes, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B.,  
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,  
 Cooke, P., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
 Fato, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,  
 Gardyna, S., Gande, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karatas, A., Kells, C., Maclean, C., MacDonald, P., Major, J.,  
 Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J.,  
 Matthews, C., McCarthy, M., Meldrum, J., Menes, L., Mihova, T.,  
 Mieng, V., Morrow, J., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,  
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
 Peterson, K., Phunkhang, P., Pierre, N., Raymond, S., Schupack, R.,  
 Seaman, S., Severy, P., Roman, J., Roy, A., Schauer, S., Schnapack, R.,  
 Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnapack, R.,  
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,  
 Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
 JOURNAL  
 COMMENT  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L2111  
 Center clone name: 20\_B\_24

FEATURES  
 source  
 1. 175020  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17"  
 /clone="RP11-20824"  
 /clone\_lib="RP11-20824 Human Male BAC"  
 complement(155..464)  
 /rpt\_family="AluSp"  
 517..547  
 /rpt\_family="TTG)n"  
 complement(559..835)  
 /rpt\_family="AluSx"  
 2222..2519  
 /rpt\_family="AluSx"  
 2849..2976  
 /rpt\_family="AluJo"  
 complement(3136..3439)  
 /rpt\_family="AluSx"  
 4309..4614  
 /rpt\_family="AluY"  
 6173..6199  
 /rpt\_family="(G)n"  
 6497..6665  
 /rpt\_family="L1MC4a"  
 complement(6667..6993)  
 /rpt\_family="AluY"  
 7077..7345  
 /rpt\_family="AluSx"  
 complement(7604..7915)

```

repeat_region /rpt_family="AluSx"
complement(17920..8220)
/rpt_family="AluSx"
repeat_region complement(8364..8662)
/rpt_family="AluY"
repeat_region 8694..8831
/rpt_family="AluUb"
repeat_region 8832..9133
/rpt_family="AluSg"
repeat_region 9143..9415
/rpt_family="AluSg"
repeat_region 10270..10279
/rpt_family="AluSg"
repeat_region 10280..10574
/rpt_family="AluSx"
repeat_region 10575..10590
/rpt_family="AluSg"
repeat_region 10591..10872
/rpt_family="AluSg"
repeat_region 10873..10886
/rpt_family="AluSg"
repeat_region 11232..11395
/rpt_family="AluSg"
repeat_region 11396..11711
/rpt_family="AluY"
repeat_region 11712..11728
/rpt_family="AluY"
repeat_region 12424..12558
/rpt_family="AluSg"
repeat_region complement(13786..14088)
/rpt_family="AluSx"
repeat_region complement(14100..14401)
/rpt_family="AluSg"
repeat_region complement(14524..14691)
/rpt_family="AluSg"
repeat_region 14827..15139
/rpt_family="AluSx"
repeat_region 15191..15283
/rpt_family="Alu"
repeat_region complement(15224..15467)
/rpt_family="AluSg"
repeat_region 15496..15781
/rpt_family="AluSg"
repeat_region 15788..15981
/rpt_family="AluSg/x"
repeat_region complement(15983..16040)
/rpt_family="AluSg"
repeat_region 16056..16220
/rpt_family="AluSg/x"
repeat_region 16330..16510
/rpt_family="Alu"
repeat_region complement(16512..16745)
/rpt_family="AluUb"
repeat_region 16746..17024
/rpt_family="Alu"
repeat_region complement(17025..17325)
/rpt_family="AluY"
repeat_region complement(17331..17625)
/rpt_family="AluUb"
repeat_region complement(17455..17459)
/rpt_family="AluSg"
repeat_region complement(17525..17538)
/rpt_family="AluSg"
repeat_region complement(17549..17588)
/rpt_family="AluSg"
repeat_region complement(17599..17607)
/rpt_family="AluSg"
repeat_region complement(17623..17627)
/rpt_family="AluSg"
repeat_region 17626..17931
/rpt_family="AluSg"
repeat_region complement(17932..18209)
/rpt_family="AluSg"

```

```

Query Match 54.7% Score 1061, DB 9, Length 175020;
Best Local Similarity 99.9%, Pred. No. 0;
Matches 1111, Conservative 0, Mismatches 1, Indels 0, Gaps 0;

QY CAGGCGCTGTGAACCTGGAAGCTGTTCCTGCGCAATGTTCTGTGAATGCTGCGCCCACTTT 866
DB CAGGCGCTGTGAACCTGGAAGCTGTTCCTGCGCAATGTTCTGTGAATGCTGCGCCCACTTT 91984
QY TCGGCCCATTTGGCTTGGCTGGTGTCCGGGCAACGAAGGGGGTCCCGACAGACCCCTGT 926
DB TCGGCCCATTTGGCTTGGCTGGTGTCCGGGCAACGAAGGGGGTCCCGACAGACCCCTGT 91924
QY ATTGTGCTCTACAGAGGGCATCGAGCCCTCTAGTGGAGATATTTGGCAACTGGCATG 986
DB ATTGTGCTCTACAGAGGGCATCGAGCCCTCTAGTGGAGATATTTGGCAACTGGCATG 91864
QY TGGAAAGAGTGTCTTCAAGCTGCGGCAACAGACCCGGGCAAGCCCATGTGGCTATGGAGCCA 1046
DB TGGAAAGAGTGTCTTCAAGCTGCGGCAACAGACCCGGGCAAGCCCATGTGGAGCCA 91863
QY GCAAGAGGCTGGCAGAGGCTGGGCTGGGGAGATGTAACCGATGAAGACCCCACT 1106
DB GCAAGAGGCTGGCAGAGGCTGGGCTGGGGAGATGTAACCGATGAAGACCCCACT 91803
QY CTGAGAGCTCAGAGGCCCCCATCTTCTTAAGACACCCCACTGAGAGACCCCACTTT 1166
DB CTGAGAGCTCAGAGGCCCCCATCTTCTTAAGACACCCCACTGAGAGACCCCACTTT 91743
QY CTCAACCTTACCCCAAGCCCTCAGAGCTCAACAGATTTGTCTAAGATGACGACCCGATTC 1226
DB CTCAACCTTACCCCAAGCCCTCAGAGCTCAACAGATTTGTCTAAGATGACGACCCGATTC 91683
QY AGGCTAAAGTGGAGCTGAGATCCAGCTCTCTAACCCTCAGAGGAGATGCTTGGCATG 1286
DB AGGCTAAAGTGGAGCTGAGATCCAGCTCTCTAACCCTCAGAGGAGATGCTTGGCATG 91633
QY GCACTTCAATGTGCTTGAAGAACTTCGAGATGTGTGAGAGGCAATGCTTGAACATGACGG 1346
DB GCACTTCAATGTGCTTGAAGAACTTCGAGATGTGTGAGAGGCAATGCTTGAACATGACGG 91563
QY GTTGTGTATCTTGAACCTCGGTGTATCTTTCTGGGGCCCAAGCTGTGCTGACATCT 1406
DB GTTGTGTATCTTGAACCTCGGTGTATCTTTCTGGGGCCCAAGCTGTGCTGACATCT 91503
QY CTTTCCGTTGAAGAAATATAGGGTATATTTCTTCTGAGAGTGAACATTAACCCCA 1466
DB CTTTCCGTTGAAGAAATATAGGGTATATTTCTTCTGAGAGTGAACATTAACCCCA 91443
QY GATGAGAGATAGGGGTATGCTAGACCTGTGCTTCTCGGAATTTGGATAGTATTTT 1526
DB GATGAGAGATAGGGGTATGCTAGACCTGTGCTTCTCGGAATTTGGATAGTATTTT 91383
QY GATGAGAGATAGGGGTATGCTAGACCTGTGCTTCTCGGAATTTGGATAGTATTTT 91324
DB GATGAGAGATAGGGGTATGCTAGACCTGTGCTTCTCGGAATTTGGATAGTATTTT 1567
QY CAGGCCCCCACTTATTTATTTCTGATCACTCTGAGAGAGAGGAGTTTGCANGT 1586
DB CAGGCCCCCACTTATTTATTTCTGATCACTCTGAGAGAGAGGAGTTTGCANGT 91333
QY GATGAGAGATAGGGGTATGCTAGACCTGTGCTTCTCGGAATTTGGATAGTATTTT 1646
DB GATGAGAGATAGGGGTATGCTAGACCTGTGCTTCTCGGAATTTGGATAGTATTTT 91263
QY GTTAATTAACCCCAATGTTAATGAAGCGGAATTAAGGCTCCGAGCTAAGAGGACTGAGCTA 1706
DB GTTAATTAACCCCAATGTTAATGAAGCGGAATTAAGGCTCCGAGCTAAGAGGACTGAGCTA 91203
QY GGGTCTCACTGATGATGAGAGAGAGGCTGGAGATCTGAACCAAGGCTGAGGCAAGGG 1766
DB GGGTCTCACTGATGATGAGAGAGAGGCTGGAGATCTGAACCAAGGCTGAGGCAAGGG 91143
QY CGAGCTGCGTAAAGATGGGTCTGAGAGTGAAGAGGAGGAGGAGCTGATGAGAGT 1826
DB CGAGCTGCGTAAAGATGGGTCTGAGAGTGAAGAGGAGGAGGAGCTGATGAGAGT 91083
QY CGAGCTGCGTAAAGATGGGTCTGAGAGTGAAGAGGAGGAGGAGCTGATGAGAGT 91024
DB CGAGCTGCGTAAAGATGGGTCTGAGAGTGAAGAGGAGGAGGAGCTGATGAGAGT

```

QY 1827 GCCCATGAGAGTAAGGGAGACGCTTCCGGGCGGATGCGAGGCTGGGCTGATCTGTATCT 1886  
 DB 91023 GCCCAATGGAGTAAGGGAGACGCTTCCGGGCGGATGCGAGGCTGGGCTGATCTGTATCT 90964  
 QY 1887 GAAGCCCTCGGATTAAGCGCGCTTGACCGCC 1918  
 DB 90963 GAAGCCCTCGGATTAAGCGCGCTTGACCGCC 90932

RESULT 4  
 AX071778 397 bp DNA linear PAT 25-JAN-2001  
 LOCUS Sequence 2250 from Patent WO0102568.  
 DEPOSITION AX071778  
 ACCESSION AX071778.1 GI:12582129  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.,  
 Kaessam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,  
 Lamson, G., Drmanac, R., Crenkajkov, R., Drmanac, S., Dickson, M.,  
 Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.  
 Human genes and gene expression products  
 Patent: WO 0102568-A 2250 11-JAN-2001;  
 CHIRON CORPORATION (US) / HYSEQ, INC. (US)

TITLE  
 JOURNAL  
 CHIRON CORPORATION (US) / HYSEQ, INC. (US)

FEATURES  
 SOURCE  
 1. 397  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

BASE COUNT 74 a 116 c 125 g 82 t

ORIGIN

Query Match 10.3%; Score 200; DB 6; Length 397;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-97;  
 Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 716 TATGTCGACATGAGTGGCTAATGTAATGTTGTCGCCGAGCTGCGCAACACACTTGAG 775  
 DB 31 TATGTCGACATGAGTGGCTAATGTAATGTTGTCGCCGAGCTGCGCAACACACTTGAG 90  
 QY 776 GCCACTGGCGTACCTGCTATGACGCCACCGAGCGCTGTGAATCGAGACTGTTCTG 835  
 DB 91 GCCACTGGCGTACCTGCTATGACGCCACCGAGCGCTGTGAATCGAGACTGTTCTG 150  
 QY 836 CGGCATGTTCTGATGAGTGGCTGCGCCACTTTTGGCCCATTTGGCTGTGCTGCTCGG 895  
 DB 151 CGGCATGTTCTGATGAGTGGCTGCGCCACTTTTGGCCCATTTGGCTGTGCTGCTCGG 210  
 QY 896 GCACCAAGAGGGGGTGCCCA 915  
 DB 211 GCACCAAGAGGGGGTGCCCA 230

RESULT 5  
 AL669840 10355 bp DNA linear ROD 05-APR-2002  
 LOCUS Mouse DNA sequence from clone RP23-43J1 on chromosome 11, complete  
 DEFINITION  
 Sequence.  
 ACCESSION AL669840  
 VERSION AL669840.5 GI:20068701  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS Bates, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,

COMMENT  
 Cambridge, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk  
 On Apr 7, 2002 this sequence version replaced gi:19309585.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: EMBL, SWISSPROT, TrEMBL, WP, WormPep, information on the WormPep  
 database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep RP23-43J1 is  
 from the RP23-43 Mouse PAC Library  
 constructed by the group of Pieter de Jong.  
 For further details see http://www.chori.org/bacpac/home.htm  
 VECTOR: pBACe3.6.

FEATURES  
 SOURCE  
 1. 103555  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="11"  
 /clone="RP23-43J1"  
 /clone\_1b="RP23-23"

BASE COUNT 27308 a 23935 c 24282 g 28026 t

ORIGIN

Query Match 2.7%; Score 52; DB 10; Length 10355;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 284 GCGCTGAGTGGCGGCGCGGAGCGCGGCTGCTGCGCGGAGCC 335  
 DB 17422 GCGCTGAGTGGCGGCGCGGAGCGCGGCTGCTGCGCGGAGCC 17473

RESULT 6  
 AC048361 204628 bp DNA linear HTG 27-MAR-2003  
 LOCUS Mus musculus chromosome 11 clone RP23-163A8 map 11, WORKING DRAFT  
 DEFINITION  
 AC048361  
 ACCESSION AC048361.4 GI:29294161  
 VERSION  
 KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS Birren, B., Nuebaum, C. and Lander, E.  
 TITLE Unpublished  
 JOURNAL  
 REFERENCE  
 AUTHORS Birren, B., Linton, L., Nuebaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barne, N., Baerlein, V., Bada, F.,  
 Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,  
 Campiano, A., Casale, A., Choepel, Y., Colangelo, M., Collins, S.,  
 Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,  
 Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
 Gagegan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
 Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
 Klein, J., Labrecque, K., Lamazares, R., Landers, T., Lehotzky, J.,  
 Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Margulis, N.,

**TITLE**  
JOURNAL  
**REFERENCE**  
AUTHORS

McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McSheeters, R.,  
Melartin, J., Meneses, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talmas, J.,  
Testaye, S., Theodore, J., Tittel, A., Travers, M., Triggillo, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wymann, D., Ye, W., J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Submitted (14-APR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 204628)

Birren, B., Nussbaum, C., Lander, E., Abouelella, A., Allen, N.,  
Anderson, S., Arachchi, H.M., Barna, N., Buetten, V., Bloom, T.,  
Boguslavsky, L., Boukhgalter, B., Camarero, J., Chang, J., Choepel, Y.,  
Collamore, A., Cook, A., Cooke, P., Corum, B., DeArillano, K.,  
Diaz, J.S., Dodge, S., Dooley, K., Dorris, J., Erickson, J., Faro, S.,  
Fitzgerald, M., Gage, D., Gage, J., Gage, J., Gage, J.,  
Graham, U., Grand-Pierre, N., Hales, N., Hagopian, D., Hago, B.,  
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R.,  
Lindblad-Toh, K., Liu, G., Liu, A., Mabbitt, R., Maclean, C.,  
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,  
Meldrum, J., Meneses, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,  
Nguyen, C., Nicol, R., Nordu, C., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, J., Peterson, K., Phunphang, P., Pierre, N.,  
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rice, C., Rogov, P.,  
Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,  
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,  
Talmas, J., Testaye, S., Theodore, J., Tittel, A., Travers, M.,  
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wymann, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

**TITLE**  
JOURNAL  
**COMMENT**

Submitted (127-MAR-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 27, 2003 this sequence version replaced gi:8076973.  
All repeats were identified using RepeatMasker:  
Smith, A.P.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

Project Information

Center project name: L7284

Center clone name: 163\_A\_8

Summary Statistics

Sequencing vector: M13; M77815, 33% of reads

Sequencing vector: Plasmid; n/a; 67% of reads

Chemistry: Dye-terminator Big Dye 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 202765 bases at least Q40

Consensus quality: 203217 bases at least Q30

Consensus quality: 203451 bases at least Q20

Insert size: 19100; agarose-ef

Insert size: 203928; sum-of-ctigs

Quality coverage: 12.1 in Q20 bases; agarose-ef

Quality coverage: 11.3 in Q20.

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 81112: contig of 81112 bp in length  
\* 81113 81112: gap of 100 bp  
\* 81213 90109: contig of 8897 bp in length  
\* 90110 90209: gap of 100 bp

**FEATURES**  
source

\* 90210 102935: contig of 12726 bp in length  
\* 102936 103035: gap of 100 bp  
\* 103036 116669: contig of 13634 bp in length  
\* 116670 116769: gap of 100 bp  
\* 116770 131956: contig of 15187 bp in length  
\* 131957 132056: gap of 100 bp  
\* 132057 148455: contig of 16399 bp in length  
\* 148456 148555: gap of 100 bp  
\* 148556 200359: contig of 51804 bp in length  
\* 200360 200459: gap of 100 bp  
\* 200460 204628: contig of 4169 bp in length.

Location/Qualifiers

1..204628

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/chromosome="11"

/map="11"

/clone="RP23-163A8"

/clone\_11b="RP21-23 Female Mouse BAC"

1..81112

/note="assembly\_fragment"

/clone\_end:SP6

vector\_side:left"

81213..90109

/note="assembly\_fragment"

90210..102935

/note="assembly\_fragment"

103036..116669

/note="assembly\_fragment"

116770..131956

/note="assembly\_fragment"

132057..148455

/note="assembly\_fragment"

148556..200359

/note="assembly\_fragment"

200460..204628

/note="assembly\_fragment"

clone\_end:TV

vector\_side:right"

BASE COUNT 54328 a 48238 c 47410 g 53947 t 705 others

ORIGIN

Query Match

Best Local Similarity 100.0%; Score 52; DB 2; Length 204628;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 168980 GCGCTGAGCTGGCGCGCGGAGCGCGCGTGTGCTGCTGCGCGAGCC 169031

284 GCGCTGAGCTGGCGCGCGGAGCGCGCGTGTGCTGCTGCGCGAGCC 335

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||



```

misc_feature      /db_xref="taxon:9606"
                  /note="1 of 2 allelic variants (6494 is other entry)"
BASE COUNT       6 a 17 c 18 g 10 t
ORIGIN
Query Match      2.6%; Score 51; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 6e-16;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 594 GCCGGAAGGATGATGCGGCTAGCGGCTGATGATGCTAGCTAGCTGCCC 644
Db 1 GCCTGAAGGATGATGCGGCTAGCGGCTGATGATGCTAGCTAGCTGCCC 51

RESULT 8
AC114124         245259 bp DNA linear HTG 13-MAY-2003
LOCUS           Rattus norvegicus clone CH230-40F15, *** SEQUENCING IN PROGRESS
DEFINITION      *** 2 unordered pieces.
ACCESSION       AC114124
VERSION         AC114124.4 GI:30579659
KEYWORDS        HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE          Rattus norvegicus (Norway rat)
ORGANISM        Rattus norvegicus
                Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                Rattus.
REFERENCE
  1 (bases 1 to 245259)
  Muzny, D., Marle, J., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J.,
  Allen, C., Allen, H., Alsebrook, S., Amin, A., Anguiano, D.,
  Ayala-Becchi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
  Baldwin, D., Bandaru, A., Barber, M., Barnstead, M., Benahmed, F.,
  Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
  Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
  Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,
  Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
  Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
  Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dedrich, D.,
  Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
  Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K.,
  Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
  Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
  Frazer, C.M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
  Georgerogiou, E., Geor, K., Gill, R., Grady, M., Guerra, M., Guevara, W.,
  Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
  Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
  Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M.,
  Hollins, B., Howell, S., Hulik, S., Hume, J., Idler, D., Jackson, A.,
  Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
  Karpach, S., Kelly, S., Kelly, S., Khan, Z., Kling, L., Kovac, C.,
  Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
  Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
  Lorenshewa, L., Loulsegue, H., Lozano, R., Lu, X., Ma, J.,
  Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
  Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, B.,
  Mathew, S., McLeod, M.P., McNeill, T.Z., Meenen, B.,
  Mlawale, S., Mlawale, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
  Morgan, M., Morris, K., Morris, S., Muidasa, M., Murphy, M., Nair, L.,
  Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
  Nsamenang, C., Okeanu, G., Olampong, A., Pal, S., Parks, K.,
  Patel, S., Paul, H., Perez, A., Perez, L., Plimack, C.,
  Plotter, F., Polidexter, A., Popovic, D., Primm, B., Pu, L.,
  Puzos, M., Quiriz, J., Rachlin, B., Reeves, K., Regier, M.A., Reigh, R.,
  Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
  Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
  Sanders, W., Savary, G., Scherer, S., Scott, S., Shatman, S., Shen, H.,
  Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smalls, D.,
  Sneed, A., Sodergren, E., Song, X.-Z., Sotelle, R., Sosa, J.,
  Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
  Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K.,
  Valas, R., Vera, V., Villaseca, D., Waldron, R., Walker, B., Wang, J.,

```

```

TITLE           JOURNAL
JOURNAL         Unpublished
REFERENCE       2 (bases 1 to 245259)
AUTHORS         Worley, K.C.
TITLE           Direct Submission
JOURNAL         Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
                of Molecular and Human Genetics, Baylor College of Medicine, One
                Baylor Plaza, Houston, TX 77030, USA
                3 (bases 1 to 245259)
                Rat Genome Sequencing Consortium.
                Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
                of Molecular and Human Genetics, Baylor College of Medicine, One
                Baylor Plaza, Houston, TX 77030, USA
                On May 13, 2003 this sequence version replaced gi:23269093.
                The sequence in this assembly is a combination of BAC based reads
                and whole genome shotgun sequencing reads assembled using Atlas
                (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
                in the feature table below represents a scaffold in the Atlas
                assembly (a 'contig-scaffold'). Within each contig-scaffold,
                individual sequence contigs are ordered and oriented, and separated
                by sized gaps filled with Ns to the estimated size. The sequence
                may extend beyond the ends of the clone and there may be sequence
                contigs within a contig-scaffold that consist entirely of whole
                genome shotgun sequence reads. Both end sequences and whole genome
                shotgun sequence only contigs will be indicated in the feature
                table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GLND
Center clone name: CH230-40F15
----- Summary Statistics
Assembly program: Atlas 3.0
Consensus quality: 195275 bases at least Q40
Consensus quality: 19584 bases at least Q30
Consensus quality: 202510 bases at least Q20
Estimated insert size: 205179; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 244085: contig of 244085 bp in length
* 244086 244185: gap of unknown length
* 244186 245259: contig of 1074 bp in length.
Location/Qualifiers
1. 245259
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-40F15"
14909..17613
/note="wgs contig"
35833..38580
/note="wgs contig"
91665..93428

```



/note="wgs\_contig"  
 mlec\_feature 113391..114894  
 BASE COUNT 53423 a 48733 c 48928 g 53896 t 40279 others  
 ORIGIN

Query Match 2.2% Score 42; DB 2; Length 245259;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-13;  
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 383 AACATGAGTCATCTTCATGCGCTTGACCTTGCGCAATCTG 424  
 Db 133948 AACATGAGTCATCTTCATGCGCTTGACCTTGCGCAATCTG 133989

RESULT 9  
 AC133243 177877 bp DNA linear HTG 10-SEP-2002  
 LOCUS Rattus norvegicus clone CH230-10014, \*\* SEQUENCING IN PROGRESS  
 DEFINITION \*\*\* 59 unordered pieces.

ACCESSION AC133243  
 VERSION AC133243.2 GI:22772764  
 KEYWORDS HTG; HTGS\_PHASE1.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus

Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE  
 1 (bases 1 to 177877)  
 Muzny, D., Marle, Metzger, M., Lee, A., Adamson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Anyalelechi, V., Aoyagi, A., Ayodeji, M., Bacc, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bernal, K., Blair, J., Blankensbury, K., Blyth, P., Brown, M., Bryant, N., Burch, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Day-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durkin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C., Gabisi, A., Gant, R., Garcia, A., Garner, T., Garza, M., Gebregorjisi, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hollins, B., Howells, S., Hulik, S., Hume, J., Idelchik, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshew, L., Louisse, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindratne, W., Mahmoud, M., Malloy, K., Mangun, A., Mangun, B., Mapua, P., Martin, K., Martin, K., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwako, O., Okunuga, G., Olarnunsgoon, A., Pal, S., Parks, K., Paternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, W., Savery, G., Scherer, S., Scott, G., Shattman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C., Smales, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steele, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, D., Tingey, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlasczyk, R.,

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

## COMMENT

Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, Y., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Welser, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 177877)  
 Rat Genome Sequencing Consortium.  
 Submitted (08-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 177877)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (10-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 On Sep 10, 2002 this sequence version replaced gi:22758715.

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 Project Information  
 Center project name: GDRJ  
 Center clone name: CH230-10014  
 Summary Statistics  
 Sequencing vector: Plasmid;  
 Chemistry: Dye-terminator Big Dye 1004 of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 120753 bases at least Q40  
 Consensus quality: 127137 bases at least Q30  
 Consensus quality: 132199 bases at least Q20

NOT: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_drafc\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_drafc_data.html)).  
 NOTE: This is a 'working draft' sequence. It currently consists of 59 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	1033	contig of 1033 bp in length
1034	1133	gap of unknown length
1134	2150	contig of 1017 bp in length
2151	2250	gap of unknown length
2251	3283	contig of 1033 bp in length
3284	3384	gap of unknown length
3384	4384	contig of 1001 bp in length
4385	4484	gap of unknown length
4485	5726	contig of 1242 bp in length
5727	5826	gap of unknown length
5827	6841	contig of 1015 bp in length
6842	6941	gap of unknown length
6942	8340	contig of 1399 bp in length
8341	8440	gap of unknown length
8441	9752	contig of 1312 bp in length
9753	9852	gap of unknown length
9853	11013	contig of 1161 bp in length
11014	11114	gap of unknown length
11114	12508	contig of 1395 bp in length
12509	12608	gap of unknown length
12609	14066	contig of 1458 bp in length
14067	14166	gap of unknown length
14167	15223	contig of 1057 bp in length
15224	15323	gap of unknown length
15324	16817	contig of 1494 bp in length
16818	16917	gap of unknown length
16918	16907	contig of 1990 bp in length
16907	19007	gap of unknown length

Query Match

Query Match	Score 31	DB 2	Length 17877
19008	20556: contig of 1549 bp in length		
20557	20656: gap of unknown length		
20657	22042: contig of 136 bp in length		
22043	22142: gap of unknown length		
22143	23173: contig of 1031 bp in length		
23174	23273: gap of unknown length		
23274	24424: contig of 1151 bp in length		
24425	24524: gap of unknown length		
24525	26395: contig of 1871 bp in length		
26396	26495: gap of unknown length		
26496	27656: contig of 1161 bp in length		
27657	27756: gap of unknown length		
27757	28988: contig of 1232 bp in length		
28989	29088: gap of unknown length		
29089	31014: contig of 1926 bp in length		
31015	31114: gap of unknown length		
31115	33214: contig of 2100 bp in length		
33215	33314: gap of unknown length		
33315	35513: contig of 2199 bp in length		
35514	35613: gap of unknown length		
35614	37742: contig of 2129 bp in length		
37743	37842: gap of unknown length		
37843	38950: contig of 1108 bp in length		
38951	39050: gap of unknown length		
39051	41722: contig of 2672 bp in length		
41723	41822: gap of unknown length		
41823	43959: contig of 2137 bp in length		
43960	44059: gap of unknown length		
44060	47505: contig of 3446 bp in length		
47506	47605: gap of unknown length		
47606	50441: contig of 2836 bp in length		
50442	50541: gap of unknown length		
50542	53541: contig of 3000 bp in length		
53542	53641: gap of unknown length		
53642	55665: contig of 2024 bp in length		
55666	55765: gap of unknown length		
55766	57645: contig of 1880 bp in length		
57646	57745: gap of unknown length		
57746	59602: contig of 1857 bp in length		
59603	59702: gap of unknown length		
59703	62147: contig of 2445 bp in length		
62148	62247: gap of unknown length		
62248	65121: contig of 2874 bp in length		
65122	65221: gap of unknown length		
65222	68430: contig of 3229 bp in length		
68431	68550: gap of unknown length		
68551	73413: contig of 4863 bp in length		
73414	73513: gap of unknown length		
73514	76578: contig of 1065 bp in length		
76579	76678: gap of unknown length		
76679	79293: contig of 2617 bp in length		
79296	79395: gap of unknown length		
79396	81954: contig of 2559 bp in length		
81955	82054: gap of unknown length		
82055	85014: contig of 2960 bp in length		
85015	85114: gap of unknown length		
85115	89088: contig of 3974 bp in length		
89089	89188: gap of unknown length		
89189	93407: contig of 4219 bp in length		
93408	93507: gap of unknown length		
93508	97240: contig of 3733 bp in length		
97241	97340: gap of unknown length		
97341	103099: contig of 5759 bp in length		
103100	103199: gap of unknown length		
103200	106048: contig of 2849 bp in length		
106049	106148: gap of unknown length		
106149	112083: contig of 5935 bp in length		
112084	112183: gap of unknown length		
112185	115595: contig of 3412 bp in length		
115596	123308: contig of 6613 bp in length		

Best Local Similarity 100.0%; Pred. No. 3,5e-05; Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1909 GTTGACCGCCCAAAAAAAAAAAAAAAAAA 1939

Db 94363 GTTGACCGCCCAAAAAAAAAAAAAAAAAA 94333

## RESULT 10

LOCUS

AB054499

413 bp DNA linear MAM 27-JUN-2001

## DEFINITION

Physeter catodon DNA, SINE flanking sequence Sp2 locus.

## ACCESSION

AB054499

## VERSION

AB054499.1 GI:114549337

## KEYWORDS

Physeter catodon (sperm whale)

## SOURCE

Physeter catodon

## ORGANISM

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Physeteridae; Physeter.

## REFERENCE

1 (bases)

## AUTHORS

Nikaido, M., Matsumo, F., Hamilton, H., Brownell, R.L., Jr., Cao, Y., Ding, W., Zuo, Y., Schedlock, A.M., Fordyce, R.E., Hasegawa, M., and Okada, N.

## TITLE

Retroposon analysis of major cetacean lineages: the monophyly of toothed whales and the paraphyly of river dolphins

## JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 98 (13), 7384-7389 (2001)

## MEDLINE

21310001

## PUBMED

11416211

## REFERENCE

2 (bases 1 to 413)

## AUTHORS

Nikaido, M.

## TITLE

Direct Submission

## JOURNAL

Submitted (17-JAN-2001) Masato Nikaido, Tokyo Institute of Technology, Department of Biological Sciences, Midori-ku

## MEDLINE

Negatsuta-cho 4259, Yokohama-shi, Kanagawa-ken 2268501, Japan

## PUBMED

(E-mail:mnikaido@bio.itech.ac.jp, Tel:81459245742, Fax:81459245835)

## FEATURES

Location/Qualifiers

## SOURCE

1. 413

## FEATURES

/organism="Physeter catodon"

## FEATURES

/mol\_type="genomic DNA"

## FEATURES

/db\_xref="taxon:9755"

## FEATURES

misc\_feature 1..413

## FEATURES

/note="SINE flanking sequence Sp2 locus"

## FEATURES

BASE COUNT 153 a 77 c 112 g 71 t

## FEATURES

ORIGIN

## FEATURES

Query Match 1.4%; Score 27; DB 4; Length 413;

## FEATURES

Best Local Similarity 100.0%; Pred. No. 0.0068; Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## FEATURES

Cy 1913 ACCGCCCAAAAAAAAAAAAAAAAAA 1939

## FEATURES

Db 310 ACCGCCCAAAAAAAAAAAAAAAAAA 336

## FEATURES

RESULT 11

## FEATURES

LOCUS

## FEATURES

G63447

## FEATURES

SHGC-141356 Human Homo sapiens STS genomic, sequence tagged site.

## FEATURES

ACCESSION

## FEATURES

G63447.1 GI:6600566

## FEATURES

VERSION

## FEATURES

G63447.1 GI:6600566

## FEATURES

KEYWORDS

## FEATURES

STS.

## FEATURES

Homo sapiens (human)

## FEATURES

Homo sapiens

## FEATURES

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## FEATURES

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

## FEATURES

REFERENCE

## FEATURES

1 (bases 1 to 561)

## FEATURES

Olivier, M. and Cox, D.R.

## FEATURES

Unpublished, Olivier, M., Cox, D.R. (2000)

## FEATURES

JOURNAL

## FEATURES

Unpublished (2000)

## FEATURES

COMMENT

## FEATURES

Contact: Michael Olivier, David R. Cox

## FEATURES

Stanford Human Genome Center  
Stanford University School of Medicine  
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA  
Tel: (650) 320-5800  
Fax: (650) 320-5801  
Email: olivier@shgc.stanford.edu  
Primer A: AATTCAGCTGTGATTCGAGC  
Primer B: TGATCCACGAGCTGAGACAGA  
STS size: 346  
PCR Profile:

Initial incubation: 95 degrees C for 10 minutes  
Denaturation: 94 degrees C for 30 seconds  
Annealing: 60 degrees C for 30 seconds  
Polymerization: 72 degrees C for 23 seconds  
PCR Cycles: 30  
Thermal Cycler: Perkin Elmer 9700  
Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
Amplifig Gold Polymerase: 0.07 units/uL  
Total Vol: 5 uL

Buffer:  
MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 8.3

Finished human sequence in NCBI. STSs designed and developed at the  
Stanford Human Genome Center.

FEATURES  
source  
1. 561  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="12"  
/clone\_lib="Human"

STS  
primer bind 208..553  
primer bind complement(531..553)

BASE COUNT 176 a 101 c 133 g 151 t  
ORIGIN

Query Match 1.4%; Score 27; DB 11; Length 561;  
Best Local Similarity 100.0%; Pred. No. 0.0067;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1913 ACCGCCAAGGAAAAAAAAAAAAA 1939  
DB 491 ACCGCCAAGGAAAAAAAAAAAAA 517

RESULT 12  
LDAJ3162/c 616 bp DNA linear PLN 12-JAN-2001  
LOCUS LDAJ3162  
DEFINITION Laminaria digitata DNA, simple tandem repeat Ld2-158.  
ACCESSION AJ003162  
VERSION AJ003162.1 GI:3392938  
KEYWORDS  
SOURCE Laminaria digitata  
ORGANISM Laminaria digitata  
Eukaryota; stramenopiles; Phaeophyceae; Laminariales;  
Laminariaceae; Laminaria.

REFERENCE  
AUTHORS Billoot C., Rousvoal S., Estoup A., Eppien J.T.,  
Saunton-Leprade P., Valero M. and Klotz G.S.  
TITLE Isolation and characterization of microsatellite markers in the  
nuclear genome of the brown alga Laminaria digitata (Phaeophyceae)  
JOURNAL Mol. Ecol. 7 (12), 1778-1780 (1998)  
MEDLINE 99076304  
PUBMED 9859207  
REFERENCE 2 (bases 1 to 616)

AUTHORS Billoot C.  
TITLE Direct Submission  
JOURNAL Submitted (27-NOV-1997) Billoot C., Biologie Cellulaire et  
Moléculaire des Algues, Station Biologique CNRS, Place Georges  
Telesier, ROSCOFF, F-29682, FRANCE

COMMENT  
FEATURES  
source  
1. 616  
Location/Qualifiers

/organism="Laminaria digitata"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:80365"  
/clone\_lib="bq2-158"

BASE COUNT 167 a 134 c 143 g 172 t  
ORIGIN

Query Match 1.4%; Score 27; DB 8; Length 616;  
Best Local Similarity 100.0%; Pred. No. 0.0066;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1913 ACCGCCAAGGAAAAAAAAAAAAA 1939  
DB 590 ACCGCCAAGGAAAAAAAAAAAAA 564

RESULT 13  
BC017773 905 bp mRNA linear PRI 06-DEC-2001  
LOCUS BC017773  
DEFINITION Homo sapiens, triggering receptor expressed on myeloid cells 1,  
clone MGC:22242 IMAGE:4692680, mRNA, complete cds.  
ACCESSION BC017773.1 GI:17389458  
VERSION BC017773.1 GI:17389458  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 905)  
Straussberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,  
USA

REMARK  
COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: CLONTECH

CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcdpax1.stanford.edu](mailto:mcdpax1.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAL Plate: 36 Row: m Column: 14  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 8924261.

FEATURES  
source  
1. 905  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="locusID:54210"  
/db\_xref="taxon:9606"  
/clone="MGC:22242 IMAGE:4692680"  
/tissue\_type="lung"  
/clone\_lib="NIH MGC\_77"  
/lab\_host="DH10B"  
/note="vector: pDNR-LIB"  
46..750

CDS

/codon\_start=1  
/product="triglycerin receptor expressed on myeloid cells  
1"  
/protein\_id="AAH1773.1"  
/db\_xref="GI:17389459"  
/translation="MRKRLMGILNMLFVSEI,RAATKLTREKYEKKEGOTLTPKCDY  
LKKFSSOKAMQIIRDEMPKTLACTERPSKSHPOVCRITLLEPHGILRYMVM  
LVEBSGLVOCYIYQPKPKPHLPDRIRIVYTKGSGTGSSENSTONYKIPPTTK  
ACCPITSPRTVOAPKSTADVSTPDSININLVNVDIIRVFNIVILLAGFLSKS  
LVFSLFAVTLFSFVP"

BASE COUNT 258 a 224 c 224 g 199 t  
Query Match 1.4%; Score 27; DB 9; Length 905;  
Best Local Similarity 100.0%; Pred. No. 0.0065;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1913 ACCGCCAATAAAAAAAAAAAAAA 1939  
870 ACCGCCAATAAAAAAAAAAAAAA 896  
RESULT 14  
LOCUS BC022657 917 bp mRNA linear ROD 16-APR-2003  
DEFINITION Mus musculus hypothetical protein LOC225847, mRNA (cDNA clone  
IMAGE:4219507), partial cds.  
ACCESSION BC022657  
VERSION BC022657.1 GI:18490481  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM

REFERENCE  
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.Y., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, D., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Schetz, T.E., Brownstein, M.J., Uedini, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullen, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, X., Gibbs, R.A.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Sheshberadaran, Y.,  
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smalinski, D.E.,  
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16099-16903 (2002)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
CNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcdopaxli@stanford.edu](mailto:mcdopaxli@stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>  
Series: IRK Plate: 43 Row: K Column: 10  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis.

FEATURES  
source  
1..917  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone\_image="4219507"  
/issue\_type="Colon, normal, 5 month old male mouse."  
/clone\_1lb="NCI\_CGAP\_C024"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1..917  
/gene="LOC225847"  
/db\_xref="locusID:225847"  
1..416  
/gene="LOC225847"  
/codon\_start=3  
/product="LOC225847 protein"  
/protein\_id="AAH22657.1"  
/db\_xref="GI:18490482"  
/db\_xref="locusID:225847"  
/translation="MRKRLMGILNMLFVSEI,RAATKLTREKYEKKEGOTLTPKCDY  
LKKFSSOKAMQIIRDEMPKTLACTERPSKSHPOVCRITLLEPHGILRYMVM  
LVEBSGLVOCYIYQPKPKPHLPDRIRIVYTKGSGTGSSENSTONYKIPPTTK  
ACCPITSPRTVOAPKSTADVSTPDSININLVNVDIIRVFNIVILLAGFLSKS  
LVFSLFAVTLFSFVP"

BASE COUNT 231 a 255 c 247 g 184 t  
Query Match 1.4%; Score 27; DB 10; Length 917;  
Best Local Similarity 100.0%; Pred. No. 0.0065;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1913 ACCGCCAATAAAAAAAAAAAAAA 1939  
868 ACCGCCAATAAAAAAAAAAAAAA 894  
RESULT 15  
LOCUS BC027481 963 bp mRNA linear PRI 08-APR-2002  
DEFINITION Homo sapiens, clone IMAGE:5107160, mRNA, partial cds.  
ACCESSION BC027481  
VERSION BC027481.1 GI:20071959  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
CNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center

Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAX Plate: 52 Row: f Column: 5  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis.

## FEATURES

## source

Location/Qualifiers  
 1..963  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5107160"  
 /issue\_type="Cervix, carcinoma"  
 /clone\_lib="NIH MGC\_12"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORTc"  
 <1..499

## CDS

/codon\_start=2  
 /product="Unknown (protein for IMAGE:5107160)"  
 /protein\_id="AAH27481.1"  
 /db\_xref="GI:20071960"  
 /translation="KDGRLMRGLRGASFGSTASALELNRLVQATQATFTTYD  
 IISPAVPSRLQDGRSARLRCGQPADIRGVAKAYDTVRGILDTQTICDVASR  
 GHQKGLTAVGCVIRQLPPTVKKFLITLTAITSLLGGRNQIVPDARKDHALKRS  
 DSAQD"

BASE COUNT 190 a 311 c 297 g 165 t  
 ORIGIN

## Query Match

Best Local Similarity 1.4%; Score 27; DB 9; Length 963;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1913 ACCGCCAAAAAAAAAAAAAAAAAAAA 1939  
 |||||  
 DB 920 ACCGCCAAAAAAAAAAAAAAAAAAAA 946

Search completed: January 31, 2004, 06:50:37  
 Job time : 7223 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2004, 04:39:02 ; Search time 126 Seconds

(without alignments)  
6752.396 Million cell updates/sec

Title: US-10-017-085A-205

Perfect score: 1939  
Sequence: 1 cgcctcgccctcgagagct.....aaaaaaaaaaaaaaaa 1939

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 10

Total number of hits satisfying chosen parameters: 239297

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match Length	ID	Description
1	1.3	1525	US-09-157-603-4	Sequence 4, Appli
2	1.3	1525	US-09-587-436-4	Sequence 4, Appli
3	1.3	1525	US-08-927-165A-4	Sequence 4, Appli
4	1.3	263	US-09-091-097-26	Sequence 26, Appli
5	1.2	38	US-09-325-554-7	Sequence 7, Appli
6	1.2	102	US-07-820-154A-10	Sequence 10, Appli
7	1.2	102	US-08-097-554A-10	Sequence 10, Appli
8	1.2	102	US-08-480-640A-10	Sequence 10, Appli
9	1.2	102	US-08-295-802-10	Sequence 10, Appli
10	1.2	102	US-08-686-968C-106	Sequence 106, App
11	1.2	102	US-08-488-237A-10	Sequence 10, Appli
12	1.2	102	US-08-375-992A-10	Sequence 10, Appli
13	1.2	102	US-08-472-679H-10	Sequence 10, Appli
14	1.2	102	PCT-US93-00344-10	Sequence 10, Appli
15	1.2	108	US-07-820-154A-10	Sequence 32, Appli
16	1.2	108	US-08-097-554A-32	Sequence 32, Appli
17	1.2	108	US-08-480-640A-32	Sequence 32, Appli
18	1.2	108	US-08-295-802-32	Sequence 32, Appli
19	1.2	108	US-08-488-237A-32	Sequence 32, Appli
20	1.2	108	US-08-375-992A-32	Sequence 32, Appli
21	1.2	108	US-08-472-679H-32	Sequence 32, Appli
22	1.2	108	PCT-US93-00324-32	Sequence 32, Appli
23	1.2	111	US-08-097-554A-78	Sequence 78, Appli
24	1.2	117	US-08-480-640A-78	Sequence 78, Appli
25	1.2	117	US-08-295-802-78	Sequence 78, Appli
26	1.2	117	US-08-686-968C-130	Sequence 130, App
27	1.2	117	US-08-686-968C-144	Sequence 144, App

c 28	24	1.2	117	3	US-08-488-237A-78	Sequence 78, Appli
c 29	24	1.2	117	4	US-08-375-992A-78	Sequence 78, Appli
c 30	24	1.2	117	4	US-08-472-679H-78	Sequence 78, Appli
c 31	24	1.2	117	4	US-08-472-679H-232	Sequence 232, App
c 32	24	1.2	117	4	US-08-472-679H-246	Sequence 246, App
c 33	24	1.2	123	3	US-08-686-968C-140	Sequence 140, App
c 34	24	1.2	123	4	US-08-472-679H-242	Sequence 242, App
c 35	24	1.2	130	6	5198345-15	Patent No. 5198345
c 36	24	1.2	177	3	US-08-686-968C-147	Sequence 147, App
c 37	24	1.2	177	4	US-08-484-575A-249	Sequence 249, App
c 38	24	1.2	178	2	US-08-484-575A-16	Sequence 16, App
c 39	24	1.2	178	3	US-08-477-459-16	Sequence 16, App
c 40	24	1.2	178	3	US-08-477-459-16	Sequence 16, App
c 41	24	1.2	178	3	US-08-486-414-16	Sequence 16, App
c 42	24	1.2	178	5	PCT-US94-01826A-16	Sequence 16, App
c 43	24	1.2	178	5	PCT-US94-02252A-16	Sequence 16, App
c 44	24	1.2	182	2	US-08-484-575A-15	Sequence 15, App
c 45	24	1.2	182	3	US-08-477-459-15	Sequence 15, App

## ALIGNMENTS

```
RESULT 1
US-09-157-603-4
Sequence 4, Application US/09157603
Patent No. 6174694
GENERAL INFORMATION:
APPLICANT: Havre, Pamela A.
APPLICANT: Rice, Michael C.
APPLICANT: Holloman, William K.
TITLE OF INVENTION: RSC2 Kinase
FILE REFERENCE: 7991-034-999
CURRENT APPLICATION NUMBER: US/09/157,603
CURRENT FILING DATE: 1998-09-21
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 1525
TYPE: DNA
ORGANISM: Mus Musculus
US-09-157-603-4

Query Match      1.3% Score 26; DB 3; Length 1525;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1914 CCGCCAAA...AAAAAAAAA 1939
Db      1459 CCGCCAAA...AAAAAAAAA 1484

RESULT 2
US-09-587-436-4
Sequence 4, Application US/09587436
Patent No. 6210916
GENERAL INFORMATION:
APPLICANT: Havre, Pamela A.
APPLICANT: Rice, Michael C.
APPLICANT: Holloman, William K.
TITLE OF INVENTION: RSC2 Kinase
FILE REFERENCE: 7991-034-999
CURRENT APPLICATION NUMBER: US/09/587,436
CURRENT FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 09/157,603
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 1525
TYPE: DNA
```

ORGANISM: Mus Musculus  
US-09-587-436-4

Query Match 1.3%; Score 26; DB 3; Length 1525;  
Best Local Similarity 100.0%; Pred. No. 0.05;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1914 CCGCCAAAAA  
1459 CCGCCAAAAA  
1484

RESULT 3  
US-08-927-165A-4  
Sequence 4, Application US/08927165A  
Patent No. 6410226

GENERAL INFORMATION:

APPLICANT: Kniec, Eric B.  
APPLICANT: Holloman, William K.

APPLICANT: Rice, Michael C.  
APPLICANT: Smith, Sheryl T.

APPLICANT: Shu, Zhigang  
TITLE OF INVENTION: Mammalian and Human Rec2  
NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kimeragen, Inc.

STREET: 300 Pleasant Run  
CITY: Newtown

STATE: PA  
COUNTRY: USA

ZIP: 18940

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

SOFTWARE: PasteSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/927,165A

FILING DATE:  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:

NAME: Hansburg, Daniel  
REGISTRATION NUMBER: 36156

REFERENCE/DOCKET NUMBER: 7991-010-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-504-4444

TELEFAX: 215-504-4545

TELEX:  
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1525 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: CDNA

US-08-927-165A-4

Query Match 1.3%; Score 26; DB 4; Length 1525;  
Best Local Similarity 100.0%; Pred. No. 0.05;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1914 CCGCCAAAAA  
1459 CCGCCAAAAA  
1484

RESULT 4  
US-09-091-097-26

Sequence 26, Application US/09091097  
Patent No. 6432407

GENERAL INFORMATION:  
APPLICANT: TAKESAKO, KAZUTOH

APPLICANT: OKADO, TAKASHI  
APPLICANT: YAGIHARA, TOMOKO

APPLICANT: KURODA, MASANOBU  
APPLICANT: ONISHI, YOSHIMI

APPLICANT: KATO, IKUNOSHIN  
APPLICANT: AKIYAMA, KAZUO

APPLICANT: YASUEDA, HIROSHI  
APPLICANT: YAMAGUCHI, HIDEYO

TITLE OF INVENTION: ANTIGENIC PROTEIN ORIGINATING IN  
TITLE OF INVENTION: MALASSEZIA

NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:

ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP  
STREET: PO BOX 747

CITY: FALLS CHURCH  
STATE: VA

COUNTRY: USA  
ZIP: 22040-0747

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/091,097

FILING DATE:  
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:  
NAME: WEINER, MARC S.

REGISTRATION NUMBER: 32,181  
REFERENCE/DOCKET NUMBER: 1422-0346P

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-205-8000

TELEFAX: 703-205-8050  
INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:  
LENGTH: 263 base pairs

TYPE: nucleic acid  
STRANDEDNESS: double

TOPOLOGY: linear  
MOLECULE TYPE: CDNA to mRNA

US-09-091-097-26

Query Match 1.3%; Score 25; DB 4; Length 263;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1915 CCGCCAAAAA  
200 CCGCCAAAAA  
224

RESULT 5  
US-09-325-554-7/c

Sequence 7, Application US/09325554  
Patent No. 6410235

GENERAL INFORMATION:  
APPLICANT: Weinand, Kurt

APPLICANT: Brand, Joachim  
TITLE OF INVENTION: DNA DETECTION BY MEANS OF A STRAND REASSOCIATION COMPLEX

FILE REFERENCE: 024420-00008  
CURRENT APPLICATION NUMBER: US/09/325,554

PRIOR FILING DATE: 1999-06-04  
PRIOR APPLICATION NUMBER: 198-24-900.4

PRIOR FILING DATE: 1998-06-04  
NUMBER OF SEQ ID NOS: 18

SOFTWARE: Patent-in version 3.1  
SEQ ID NO: 7

LENGTH: 38  
TYPE: DNA

ORGANISM: Mycobacterium tuberculosis



FEATURE:  
NAME/KEY: misc signal  
LOCATION: (1) .. (1)  
OTHER INFORMATION: Phosphate linked to biotin via AminoLinker  
US-09-325-554-7

Query Match 1.2%; Score 24; DB 4; Length 38;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1916 GCCAAGAAAAAAGAAAAA 1939  
Db 25 GCCAAGAAAAAAGAAAAA 2

RESULT 6  
US-07-820-154A-10/C  
Sequence 10, Application US/07820154A  
Patent No. 5382425

GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820,154A  
FILING DATE: 19920113  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
TELEX: 422523

INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmid  
IMMEDIATE SOURCE:  
CLONE: 520-17.5 (Junction B)

FEATURE:  
NAME/KEY: CDS  
LOCATION: 85..99  
OTHER INFORMATION: /codon start= 85  
OTHER INFORMATION: /function= "translational start of hybrid protein"  
OTHER INFORMATION: /product= "N-terminal peptide"  
OTHER INFORMATION: /number= 1  
OTHER INFORMATION: /standard\_name= "Translation of synthetic DNA  
sequence"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 100..102  
IDENTIFICATION METHOD: experimental  
OTHER INFORMATION: /partial

OTHER INFORMATION: /codon start= 100  
OTHER INFORMATION: /function= "marker enzyme"  
OTHER INFORMATION: /product= "Beta-Galactosidase"  
OTHER INFORMATION: /evidence= EXPERIMENTAL  
OTHER INFORMATION: /gene= "lacZ"  
OTHER INFORMATION: /number= 2  
OTHER INFORMATION: /citation= ([1])  
PUBLICATION INFORMATION:  
AUTHORS: Ferrari, Franco A  
AUTHORS: Trach, Kathleen  
AUTHORS: Hoch, James A  
TITLE: Sequence Analysis of the spoB Locus Reveals  
TITLE: a Polycistronic Transcription Unit  
JOURNAL: J. Bacteriol.  
VOLUME: 161  
ISSUE: 2  
PAGES: 556-562  
DATE: Feb.-1985  
US-07-820-154A-10

Query Match 1.2%; Score 24; DB 1; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1916 GCCAAGAAAAAAGAAAAA 1939  
Db 78 GCCAAGAAAAAAGAAAAA 55

RESULT 7  
US-08-097-554A-10/C  
Sequence 10, Application US/08097554A  
Patent No. 5869312

GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David E  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 112  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/097,554A  
FILING DATE: July 22, 1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
TELEX: 422523

INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmid  
IMMEDIATE SOURCE:  
CLONE: 520-17.5 (Junction B)

FEATURE:  
NAME/KEY: CDS  
LOCATION: 85...99 /codon\_start= 95  
OTHER INFORMATION: /function= "Translational start of hybrid protein"  
OTHER INFORMATION: /product= "N-terminal peptide"  
OTHER INFORMATION: /number= 1  
OTHER INFORMATION: /standard\_name= "Translation of synthetic DNA  
OTHER INFORMATION: sequence"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 100..102  
IDENTIFICATION METHOD: experimental  
OTHER INFORMATION: /partial  
OTHER INFORMATION: /codon\_start= 100  
OTHER INFORMATION: /function= "marker enzyme"  
OTHER INFORMATION: /product= "Beta-Galactosidase"  
OTHER INFORMATION: /evidence= EXPERIMENTAL  
OTHER INFORMATION: /gene= "lacZ"  
OTHER INFORMATION: /number= 2  
OTHER INFORMATION: /citation= (11)  
PUBLICATION INFORMATION:  
AUTHORS: Ferrari, Franco A  
AUTHORS: Trach, Kathleen  
TITLES: Sequence Analysis of the spo0B locus Reveals  
TITLE: a Polycistronic Transcription Unit  
JOURNAL: J. Bacteriol.  
VOLUME: 161  
ISSUE: 2  
PAGES: 556-562  
DATE: Feb.-1985  
US-08-097-554A-10

Query Match 1.2%; Score 24; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939  
Db 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

RESULT 8  
US-08-480-640A-10/c  
Sequence 10, Application US/08480640A  
Patent No. 6033904  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
APPLICANT: Junker, David B.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,640A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmid  
IMMEDIATE SOURCE:  
CLONE: 520-17.5 (Junction B)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 85..99  
OTHER INFORMATION: /codon\_start= 85  
OTHER INFORMATION: /function= "Translational start of hybrid protein"  
OTHER INFORMATION: /product= "N-terminal peptide"  
OTHER INFORMATION: /number= 1  
OTHER INFORMATION: /standard\_name= "Translation of synthetic DNA  
OTHER INFORMATION: sequence"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 100..102  
IDENTIFICATION METHOD: experimental  
OTHER INFORMATION: /partial  
OTHER INFORMATION: /codon\_start= 100  
OTHER INFORMATION: /function= "marker enzyme"  
OTHER INFORMATION: /product= "Beta-Galactosidase"  
OTHER INFORMATION: /evidence= EXPERIMENTAL  
OTHER INFORMATION: /gene= "lacZ"  
OTHER INFORMATION: /number= 2  
OTHER INFORMATION: /citation= (11)  
PUBLICATION INFORMATION:  
AUTHORS: Ferrari, Franco A  
AUTHORS: Trach, Kathleen  
TITLES: Sequence Analysis of the spo0B locus Reveals  
TITLE: a Polycistronic Transcription Unit  
JOURNAL: J. Bacteriol.  
VOLUME: 161  
ISSUE: 2  
PAGES: 556-562  
DATE: Feb.-1985  
US-08-480-640A-10

Query Match 1.2%; Score 24; DB 3; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939  
Db 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

RESULT 9  
US-08-295-802-10/c  
Sequence 10, Application US/08295802  
Patent No. 6127163  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D  
APPLICANT: Junker M.S., David B  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 188  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/295,802
FILING DATE: Herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)977-9550
TELEFAX: (212)664-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmid
IMMEDIATE SOURCE:
CLONE: 520-17.5 (Junction B)
FEATURE:
NAME/KEY: CDS
LOCATION: 85..99
OTHER INFORMATION: /codon_start= 85
OTHER INFORMATION: /function= "Translational start of hybrid protein"
OTHER INFORMATION: /product= "N-terminal peptide"
OTHER INFORMATION: /number= 1
OTHER INFORMATION: /standard_name= "Translation of synthetic DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 100..102
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /partial
OTHER INFORMATION: /codon_start= 100
OTHER INFORMATION: /function= "marker enzyme"
OTHER INFORMATION: /product= "Beta-Galactosidase"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "lacZ"
OTHER INFORMATION: /number= 2
OTHER INFORMATION: /citations= (11)
PUBLICATION INFORMATION:
AUTHORS: Ferrari, Franco A
AUTHORS: Trach, Kathleen
TITLE: Sequence Analysis of the spoOB Locus Reveals
TITLE: a Polycistronic Transcription Unit
JOURNAL: J. Bacteriol.
VOLUME: 161
ISSUE: 2
PAGES: 556-562
DATE: Feb.-1985
US-08-295-802-10

```

```

Query Match 1.2%; Score 24; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939
DB 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

```

```

RESULT 10
US-08-686-968C-106/c

```

```

Sequence 106, Application US/08686968C
Patent No. 6221351
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David B.
TITLE OF INVENTION: Recombinant Swinepox Virus
FILE REFERENCE: 39119-H/JWL
CURRENT APPLICATION NUMBER: US/08/686,968C
CURRENT FILING DATE: 1996-07-25
NUMBER OF SEQ ID NOS: 231
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 106
LENGTH: 102
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Homology
US-08-686-968C-106

```

```

Query Match 1.2%; Score 24; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939
DB 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

```

```

RESULT 11
US-08-488-237A-10/c
Sequence 10, Application US/08488237A
Patent No. 6251403
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Junker, David B.
TITLE OF INVENTION: Recombinant Swinepox Virus
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESSES:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,237A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 102 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Plasmid
IMMEDIATE SOURCE:
CLONE: 520-17.5 (Junction B)

```

FEATURE:  
NAME/KEY: CDS  
LOCATION: 85..99  
OTHER INFORMATION: /codon\_start= 85  
OTHER INFORMATION: /function= "Translational start of hybrid protein"  
OTHER INFORMATION: /product= "N-terminal peptide"  
OTHER INFORMATION: /number= 1  
OTHER INFORMATION: /standard\_name= "Translation of synthetic DNA"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 100..102  
IDENTIFICATION METHOD: experimental  
OTHER INFORMATION: /partial  
OTHER INFORMATION: /codon\_start= 100  
OTHER INFORMATION: /function= "marker enzyme"  
OTHER INFORMATION: /product= "Beta-galactosidase"  
OTHER INFORMATION: /evidence= EXPERIMENTAL  
OTHER INFORMATION: /gene= "lacZ"  
OTHER INFORMATION: /number= 2  
OTHER INFORMATION: /citation= (11)  
PUBLICATION INFORMATION:  
AUTHORS: Ferrari, Franco A  
AUTHORS: Trach, Kathleen  
AUTHORS: Hoch, James A  
TITLE: Sequence Analysis of the spoOB locus Reveals  
TITLE: a Polyclonistic Transcription Unit  
JOURNAL: J. Bacteriol.  
VOLUME: 161  
ISSUE: 2  
PAGES: 556-562  
DATE: Feb.-1985  
US-08-488-737A-10

Query Match 1.2%; Score 24; DB 3; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939  
Db 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

RESULT 12  
US-08-375-992A-10/c  
Sequence 10, Application US/08375992A  
Patent No. 6328975  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
APPLICANT: Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 220  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/375,992A  
FILING DATE: Herewith  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmid  
IMMEDIATE SOURCE:  
CLONE: 520-17.5 (Junction B)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 85..99  
OTHER INFORMATION: /codon\_start= 85  
OTHER INFORMATION: /function= "Translational start of hybrid protein"  
OTHER INFORMATION: /product= "N-terminal peptide"  
OTHER INFORMATION: /number= 1  
OTHER INFORMATION: /standard\_name= "Translation of synthetic DNA"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 100..102  
IDENTIFICATION METHOD: experimental  
OTHER INFORMATION: /partial  
OTHER INFORMATION: /codon\_start= 100  
OTHER INFORMATION: /function= "marker enzyme"  
OTHER INFORMATION: /product= "Beta-Galactosidase"  
OTHER INFORMATION: /evidence= EXPERIMENTAL  
OTHER INFORMATION: /gene= "lacZ"  
OTHER INFORMATION: /number= 2  
OTHER INFORMATION: /citation= (11)  
PUBLICATION INFORMATION:  
AUTHORS: Ferrari, Franco A  
AUTHORS: Trach, Kathleen  
AUTHORS: Hoch, James A  
TITLE: Sequence Analysis of the spoOB locus Reveals  
TITLE: a Polyclonistic Transcription Unit  
JOURNAL: J. Bacteriol.  
VOLUME: 161  
ISSUE: 2  
PAGES: 556-562  
DATE: Feb.-1985  
US-08-375-992A-10

Query Match 1.2%; Score 24; DB 4; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939  
Db 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

RESULT 13  
US-08-472-679H-10/c  
Sequence 10, Application US/08472679H  
Patent No. 6497882  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
APPLICANT: Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 267  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pamela G. Salkeid  
STREET: 2000 Gallopington Hill Road  
CITY: Kentilworth  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07033

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,679H  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Salkeid, Pamela G.  
REGISTRATION NUMBER: 38,607  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 298-2135  
TELEFAX: (908) 298-5388  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmid  
IMMEDIATE SOURCE:  
CLONE: 520-17.5 (Junction B)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 85..99  
OTHER INFORMATION: /codon\_start= 85/function= "Translational start of hybrid protein"  
/product= "N-terminal peptide"/number= 1  
/standard\_name= "translation of synthetic DNA sequence"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 100..102  
IDENTIFICATION METHOD: experimental  
OTHER INFORMATION: /partial /codon\_start= 100  
/function= "marker enzyme"/product= "Beta-Galactosidase"  
/evidence= EXPERIMENTAL  
/gene= "lacZ"  
/number= 2  
/citation= ([1])  
PUBLICATION INFORMATION:  
AUTHORS: Ferrari, Franco A, Trach, Kathleen, Hoch, James A  
TITLE: Sequence Analysis of the spoOB Locus Reveals a Polycistronic Transcription Unit  
PATENT NO. 6497882  
JOURNAL: J. Bacteriol.  
VOLUME: 161  
ISSUE: 2  
PAGES: 556-562  
DATE: Feb.-1985  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-08-472-679H-10  
Query Match 1.2%; Score 24; DB 4; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/00324  
FILING DATE: 19930113  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmid  
IMMEDIATE SOURCE:  
CLONE: 520-17.5 (Junction B)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 85..99  
OTHER INFORMATION: /codon\_start= 85  
/function= "Translational start of hybrid protein"  
/product= "N-terminal peptide"  
/number= 1  
/standard\_name= "translation of synthetic DNA sequence"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 100..102  
IDENTIFICATION METHOD: experimental  
OTHER INFORMATION: /partial  
OTHER INFORMATION: /codon\_start= 100  
OTHER INFORMATION: /function= "marker enzyme"  
OTHER INFORMATION: /product= "Beta-Galactosidase"  
OTHER INFORMATION: /evidence= EXPERIMENTAL  
OTHER INFORMATION: /gene= "lacZ"  
OTHER INFORMATION: /number= 2  
OTHER INFORMATION: /citation= ([1])  
PUBLICATION INFORMATION:  
AUTHORS: Ferrari, Franco A  
AUTHORS: Trach, Kathleen  
AUTHORS: Hoch, James A  
TITLE: Sequence Analysis of the spoOB Locus Reveals a Polycistronic Transcription Unit  
JOURNAL: J. Bacteriol.  
VOLUME: 161  
ISSUE: 2  
PAGES: 556-562  
DATE: Feb.-1985  
PCT-US93-00324-10  
Query Match 1.2%; Score 24; DB 5; Length 102;  
Best Local Similarity 100.0%; Pred. No. 0.47;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939  
Db 78 GCCAAAAAAAAAAAAAAAAAAAAA 55

RESULT 15  
US-07-820-154A-32/c  
Sequence 32, Application US/07820154A  
Patent No. 5382425  
GENERAL INFORMATION:  
APPLICANT: Cochran Ph.D., Mark D.  
APPLICANT: Junker M.S., David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/820,154A  
FILING DATE: 19920113  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 108 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmid  
IMMEDIATE SOURCE:  
CLONE: 538-46.26 (Junction B)  
FEATURE:  
NAME/KEY: exon  
LOCATION: 88..102  
OTHER INFORMATION: /codon\_start= 88  
OTHER INFORMATION: /function= "Translational start of hybrid protein"  
OTHER INFORMATION: /product= "N-terminal peptide"  
OTHER INFORMATION: /number= 1  
OTHER INFORMATION: /standard\_name= "Translation of synthetic DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 103..108  
IDENTIFICATION METHOD: experimental  
OTHER INFORMATION: /partial  
OTHER INFORMATION: /codon\_start= 103  
OTHER INFORMATION: /product= "NDV Haemagglutinin-Neuraminidase"  
OTHER INFORMATION: /evidence= EXPERIMENTAL  
OTHER INFORMATION: /gene= "HN"  
OTHER INFORMATION: /number= 2  
US-07-820-154A-32

Query Match

1.2% Score 24; DB 1; Length 108;

Best Local Similarity 100.0%; Pred. No. 0.47;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1916 GCCAAAAAAAAAAAAAAAAAAAAA 1939  
Db 81 GCCAAAAAAAAAAAAAAAAAAAAA 58

Search completed: January 31, 2004, 08:03:22  
Job time: 128 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2004, 04:06:41, Search time 550 Seconds

(without alignments)  
9516.746 Million cell updates/sec

Title: US-10-017-085A-205

Perfect score: 1939  
Sequence: 1 cgcctcgcgccttcgagagct.....aaaaaaaaaaaaaaaa 1939

Scoring table: OLIGO NUC

Gapop 60.0, Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size: 10

Total number of hits satisfying chosen parameters: 1552787

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: N\_Geneseq\_19Jun03.\*

```
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1939	100.0	1939	20	AAZ34065 Human PRO853 nucle
2	1939	100.0	1939	20	AAZ34065 Human PRO853 nucle
3	1939	100.0	1939	21	AAZ34065 Human PRO853 nucle
4	1939	100.0	1939	21	AAZ34065 Human PRO853 nucle
5	1939	100.0	1939	25	ABX92437 Human PRO853 nucle
6	1803	93.0	1887	24	ABL90069 Human polynucleoti
7	1497	77.2	1779	22	ABD09338 Human drug metabol
8	779	40.2	1098	25	ABZ68100 Human secreted pro

9	698	36.0	1140	22	AAZ26578	DNA encoding human
10	698	36.0	1140	25	ABZ67229	Human secreted pro
11	396	20.4	598	21	AAZ6397	Human ORFX ORF1952
12	200	10.3	397	22	AAZ66494	Novel human polyu
13	92	4.7	186	16	AAZ19666	Human gene signatu
14	88	4.5	175	24	ABK64203	Human benign prost
15	51	2.6	51	22	AAZ19552	Human silent SNP c
16	45	2.3	45	20	AAZ34068	Human PRO853 hybr
17	45	2.3	45	21	AAZ78735	Human PRO853 hybr
18	45	2.3	45	21	AAZ46935	Probe used to scre
19	45	2.3	45	25	ABX92440	Human PRO DNA prob
20	27	1.4	1481	21	AAZ69406	Human secreted pro
21	27	1.4	2093	21	AAZ74484	Human ORFX ORF3039
22	26	1.3	159	23	ABV34595	Human prostate exp
23	26	1.3	159	23	ABV34595	Human prostate exp
24	26	1.3	179	23	ABV34595	Human prostate exp
25	26	1.3	207	24	ABZ65459	Human prostate exp
26	26	1.3	282	22	AAZ10259	Lung cancer relate
27	26	1.3	295	23	ABV13479	Human breast cance
28	26	1.3	356	22	AAZ72392	Human prostate exp
29	26	1.3	359	22	AAZ69507	Human cervical can
30	26	1.3	360	22	AAZ70842	Human cervical can
31	26	1.3	407	22	AAZ64733	Human polynucleoti
32	26	1.3	419	23	ABV61898	Human prostate exp
33	26	1.3	481	22	AAZ18861	Human breast cance
34	26	1.3	495	22	ABZ18861	Human breast cance
35	26	1.3	495	22	ABZ18861	Human foetal liver
36	26	1.3	495	22	AAZ0154	Human brain expres
37	26	1.3	495	22	AAZ25594	Human bone marrow
38	26	1.3	495	22	AAZ10219	Probe #152 for gen
39	26	1.3	495	22	AAZ13468	Probe #154 used to
40	26	1.3	495	22	AAZ100158	Probe #149 used to
41	26	1.3	495	23	ABZ25170	Human liver single
42	26	1.3	495	24	ABZ00163	Human genome-deriv
43	26	1.3	570	24	ABZ043565	Oligonucleotide fo
44	26	1.3	570	24	ABZ043567	Oligonucleotide fo
45	26	1.3	631	20	AAZ37428	Human secreted pro

#### ALIGNMENTS

RESULT 1	
ID	AAZ34065 standard; cDNA; 1939 BP.
AC	AAZ34065;
XX	
DT	07-DEC-1999 (first entry)
XX	
DE	Human PRO853 nucleotide sequence.
XX	
KW	Human; PRO: EST: expressed sequence tag; PCR primer; hybridization;
KW	probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW	secreted protein; transmembrane protein; ss.
XX	
OS	Homo sapiens.
XX	
PN	W09946281-A2.
XX	
PD	16-SEP-1999.
XX	
PF	08-MAR-1999;
XX	
PR	10-MAR-1998;
PR	11-MAR-1998;
PR	11-MAR-1998;
PR	11-MAR-1998;
PR	12-MAR-1998;
PR	13-MAR-1998;
PR	17-MAR-1998;
PR	20-MAR-1998;
PR	20-MAR-1998;
PR	98US-0078886;
PR	98US-0078910;



PR 20-MAR-1998; 98US-0078936.  
 PR 20-MAR-1998; 98US-0078939.  
 PR 25-MAR-1998; 98US-0079294.  
 PR 26-MAR-1998; 98US-0079656.  
 PR 27-MAR-1998; 98US-0079663.  
 PR 27-MAR-1998; 98US-0079664.  
 PR 27-MAR-1998; 98US-0079689.  
 PR 27-MAR-1998; 98US-0079728.  
 PR 27-MAR-1998; 98US-0079786.  
 PR 30-MAR-1998; 98US-0079920.  
 PR 30-MAR-1998; 98US-0079923.  
 PR 31-MAR-1998; 98US-0080105.  
 PR 31-MAR-1998; 98US-0080107.  
 PR 31-MAR-1998; 98US-0080165.  
 PR 31-MAR-1998; 98US-0080194.  
 PR 01-APR-1998; 98US-0080327.  
 PR 01-APR-1998; 98US-0080328.  
 PR 01-APR-1998; 98US-0080333.  
 PR 01-APR-1998; 98US-0080334.  
 PR 08-APR-1998; 98US-0081049.  
 PR 08-APR-1998; 98US-0081070.  
 PR 08-APR-1998; 98US-0081071.  
 PR 09-APR-1998; 98US-0081195.  
 PR 09-APR-1998; 98US-0081203.  
 PR 09-APR-1998; 98US-0081229.  
 PR 15-APR-1998; 98US-0081817.  
 PR 15-APR-1998; 98US-0081838.  
 PR 15-APR-1998; 98US-0081952.  
 PR 15-APR-1998; 98US-0081955.  
 PR 21-APR-1998; 98US-0082568.  
 PR 21-APR-1998; 98US-0082569.  
 PR 22-APR-1998; 98US-0082700.  
 PR 22-APR-1998; 98US-0082704.  
 PR 22-APR-1998; 98US-0082804.  
 PR 23-APR-1998; 98US-0082767.  
 PR 23-APR-1998; 98US-0082796.  
 PR 27-APR-1998; 98US-0083336.  
 PR 28-APR-1998; 98US-0083342.  
 PR 29-APR-1998; 98US-0083392.  
 PR 29-APR-1998; 98US-0083495.  
 PR 29-APR-1998; 98US-0083496.  
 PR 29-APR-1998; 98US-0083499.  
 PR 29-APR-1998; 98US-0083500.  
 PR 29-APR-1998; 98US-0083545.  
 PR 29-APR-1998; 98US-0083554.  
 PR 29-APR-1998; 98US-0083558.  
 PR 29-APR-1998; 98US-0083559.  
 PR 30-APR-1998; 98US-0083742.  
 PR 05-MAY-1998; 98US-0084356.  
 PR 06-MAY-1998; 98US-0084414.  
 PR 06-MAY-1998; 98US-0084441.  
 PR 07-MAY-1998; 98US-0084598.  
 PR 07-MAY-1998; 98US-0084600.  
 PR 07-MAY-1998; 98US-0084627.  
 PR 07-MAY-1998; 98US-0084637.  
 PR 07-MAY-1998; 98US-0084639.  
 PR 07-MAY-1998; 98US-0084640.  
 PR 07-MAY-1998; 98US-0084643.  
 PR 13-MAY-1998; 98US-0085233.  
 PR 13-MAY-1998; 98US-0085338.  
 PR 13-MAY-1998; 98US-0085339.  
 PR 13-MAY-1998; 98US-0085573.  
 PR 15-MAY-1998; 98US-0085579.  
 PR 15-MAY-1998; 98US-0085580.  
 PR 15-MAY-1998; 98US-0085582.  
 PR 15-MAY-1998; 98US-0085589.  
 PR 15-MAY-1998; 98US-0085597.  
 PR 15-MAY-1998; 98US-0085700.  
 PR 15-MAY-1998; 98US-0085704.  
 PR 18-MAY-1998; 98US-0086023.  
 PR 22-MAY-1998; 98US-0086392.  
 PR 22-MAY-1998; 98US-0086414.  
 PR 22-MAY-1998; 98US-0086430.

PR 22-MAY-1998; 98US-0086486.  
 PR 28-MAY-1998; 98US-0087098.  
 PR 28-MAY-1998; 98US-0087106.  
 PR 28-MAY-1998; 98US-0087208.  
 PR 30-JUL-1998; 98US-0094651.  
 PR 11-SEP-1998; 98US-0100038.  
 XX (GETH ) GENENTECH INC.  
 PA  
 XX  
 PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J,  
 XX  
 XX WPI, 1999-551358/46.  
 DR P-PSDB; AA141715.  
 XX  
 XX New secreted and transmembrane polypeptides and their polynucleotides,  
 PT useful for treating blood coagulation disorders, cancers and cellular  
 PT adhesion disorders -  
 PS  
 PS Claim 2, Fig 74, 530pp, English.  
 XX  
 XX The present invention describes secreted and transmembrane polypeptides  
 CC and their polynucleotides. The nucleotide sequences are useful as  
 CC sources of probes, primers, for chromosome mapping, and for generation  
 CC of antisense sequences. They can also be used to create transgenic  
 CC animals. The proteins can be used to treat a variety of diseases and  
 CC disorders, depending on their function. Diseases that may be treated  
 CC include blood coagulation disorders, cancers and cellular adhesion  
 CC disorders. They may also be used to raise antibodies. AA233891 to  
 CC AA233338, and AA141685 to AA141774 represent polynucleotide and  
 CC polypeptide sequence given in the exemplification of the present  
 CC invention.  
 XX  
 XX Sequence 1939 BP; 367 A; 549 C; 624 G; 399 T; 0 other;  
 SQ  
 Query Match 100.0%; Score 1939; DB 20; Length 1939;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1939; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGGCTCCGCTTCGAGGCTGACGCGCCGCGGCGCTTCAGAGCCTGTGACGAGCGGAT 60  
 DB 1 CGGCTCCGCTTCGAGGCTGACGCGCCGCGGCGCTTCAGAGCCTGTGACGAGCGGAT 60  
 QY 61 CGGACGCCCTGTGCGGCGCATTCAGGCGGTGCGGCGCTTCGAGGCGGAGCGCGC 120  
 DB 61 CGGACGCCCTGTGCGGCGCATTCAGGCGGTGCGGCGCTTCGAGGCGGAGCGCGC 120  
 QY 121 GCGCGGCGATGAGGCGCTGTGCGGCGCGGCGGTGCTGTGCGGCGCTTACGTGCTGT 180  
 DB 121 GCGCGGCGATGAGGCGCTGTGCGGCGCGGCGGTGCTGTGCGGCGCTTACGTGCTGT 180  
 QY 181 CTACTACAACTGTGTGAGAGCCCGCGTGTGCGGCGCATGTGCGGCGGCGCAC 240  
 DB 181 CTACTACAACTGTGTGAGAGCCCGCGTGTGCGGCGCATGTGCGGCGGCGCAC 240  
 QY 241 GCGCGTGTGACGCGGCGCGCAACAGCGGCATCGGAAGAATGACGCGCGTGTGAGCG 300  
 DB 241 GCGCGTGTGACGCGGCGCGCAACAGCGGCATCGGAAGAATGACGCGCGTGTGAGCG 300  
 QY 301 GCGCGGAGGCGGCGGTGTGCTGTGCGGCGCGGCGGAGCGGCGGAGCGGCGCTGCTT 360  
 DB 301 GCGCGGAGGCGGCGGTGTGCTGTGCGGCGCGGCGGAGCGGCGGAGCGGCGCTGCTT 360  
 QY 361 CGACCTCCGCGAGAGAGTGTGGAACATGAGTCACTTCATGAGCTTGTGAGCTTGTGCGAG 420  
 DB 361 CGACCTCCGCGAGAGAGTGTGGAACATGAGTCACTTCATGAGCTTGTGAGCTTGTGCGAG 420  
 QY 421 TCTGCGCTGTGCGGCGCTTGTGCACTGTGCTTGTGAGCTTGTGAGCGGTTGAGCAT 480  
 DB 421 TCTGCGCTGTGCGGCGCTTGTGCACTGTGCTTGTGAGCTTGTGAGCGGTTGAGCAT 480  
 QY 481 CCGTATCCCAATGCGCGGTATGAGTCTGTGCGGCGGAGCCCGGAGCGGTTAACTGTCT 540  
 DB 481 CCGTATCCCAATGCGCGGTATGAGTCTGTGCGGCGGAGCCCGGAGCGGTTAACTGTCT 540

[illegible]

QY	1621	TCCCTTGCAACCGCTACGTAGAGTAGTAAATATACCCCATGTATTATGAAGCGAATTA	1680
Db	1621	TCCCTTGCAACCGCTACGTAGAGTAGTAAATTAACCCCATGTATTATGAAGCGAATTA	1680
QY	1681	GGCTCCCGAGCTAAGGAGCACTCGCCTAGGAGTCTCACAGTGAAGAGAGGCGCTGGGAT	1740
Db	1681	GGCTCCCGAGCTAAGGAGCACTCGCCTAGGAGTCTCACAGTGAAGAGAGGCGCTGGGAT	1740
QY	1741	CTGAACCCCAAGGGTCTGAGGCGAGGGGCGAGCTGCGTAAAGTGGGTCTGAGAAAGTAGT	1800
Db	1741	CTGAACCCCAAGGGTCTGAGGCGAGGGGCGAGCTGCGTAAAGTGGGTCTGAGAAAGTAGT	1800
QY	1801	CAGGGCAGGGCAGCTGTGATATGAGGTGCCCATGGAGTAGAAGGGAGCGCCTTCGGGCGG	1860
Db	1801	CAGGGCAGGGCAGCTGTGATATGAGGTGCCCATGGAGTAGAAGGGAGCGCCTTCGGGCGG	1860
QY	1861	ATGCAAGGCTGGGGTCAATCTGTATCTGAAGCCCTTCGGAATTAAGCGCTTGACACCGCA	1920
Db	1861	ATGCAAGGCTGGGGTCAATCTGTATCTGAAGCCCTTCGGAATTAAGCGCGTTGACACCGCA	1920
QY	1921	AAAAAAAAAAAAAAAAAAAA 1939	
Db	1921	AAAAAAAAAAAAAAAAAAAA 1939	

ID	AAx87266 standard; cDNA, 1939 BP.
XX	AAx87266
XX	AAx87266;
AC	27-SEP-1999 (first entry)
DT	cDNA clone encoding human PRO853, amplified in tumour cells.
XX	
XX	PRO853; UNQ419; tumour; cancer; diagnosis; therapy; human; ds.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	128..1261
FT	/*tag= a
FT	sig_peptide
FT	128..175
FT	/*tag= b
FT	mat_peptide
FT	175..1258
XX	/*tag= c
XX	
PN	W09935170-A2.
XX	
XX	
PD	15-JUL-1999.
XX	
PF	05-JAN-1999; 99NO-US00106.
XX	
XX	
PR	20-NOV-1998; 98US-0109304.
PR	05-JAN-1998; 98US-0070440.
PR	29-APR-1998; 98US-0083500.
PR	22-MAY-1998; 98US-0086614.
PR	10-JUN-1998; 98US-0088742.
PR	10-NOV-1998; 98US-0107783.
XX	
PA	(GETH ) GENENTECH INC.
PI	Borstein D, Goddard A, Gurney AL, Hillan KJ, Lawrence DA,
PI	Roy MA, Wood WT,
XX	
DR	WPI; 1999-430385/36.
DR	P-PSDB; AAY06489.
XX	
PT	Antibody against proteins expressed in neoplastic cells, useful for
PT	tumour diagnosis and treatment
XX	
PS	Example 1; Fig 25A-B; 162pp; English.
XX	

CC This is the nucleotide sequence of cDNA clone DNA4827 (ATCC 209812)  
CC coding for human PRO83 (ONQ419) (see AY06489). The clone was  
CC isolated from a human foetal kidney tissue library. Amplification  
CC of DNA4827 (chromosome 17) occurs in various tumours, especially  
CC colon tumours, suggesting an association with tumour formation or  
CC growth. Antagonists (e.g. antibodies) directed to PRO83 may have  
CC use in cancer therapy. The invention identifies 14 genes (see  
CC AX87254-67) that are amplified in the genome of tumour cells. Such  
CC amplification is expected to be associated with overexpression of  
CC the gene product and to contribute to tumorigenesis. The encoded  
CC proteins (see AY06477-90) may be useful targets for the diagnosis  
CC and/or treatment (including prevention) of certain cancers, and may  
CC act as predictors of the prognosis of tumour treatment.

XX Sequence 1939 BP; 367 A; 549 C; 624 G; 399 T; 0 other;

Query Match 100.0%; Score 1939; DB 20; Length 1939;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1939; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTCCGCTTGGAGAGCTGACGCGCCGCGCGCTTCCAGGCTGTGACAGGCGGAT 60  
DB 1 CGCTCCGCTTGGAGAGCTGACGCGCCGCGCGCTTCCAGGCTGTGACAGGCGGAT 60  
QY 61 CGGAGCGCGCTGGCGGCGATCCAGGCGGCTGGCGGCGGAGCGCGGAGCGCGC 120  
DB 61 CGGAGCGCGCTGGCGGCGATCCAGGCGGCTGGCGGCGGAGCGCGGAGCGCGC 120  
QY 121 GCGCGGATGAGAGCGCTGCTGCTGGCGCGGCGGCTGCTGCTGCGGCTTACGCTTGT 180  
DB 121 GCGCGGATGAGAGCGCTGCTGCTGGCGCGGCGGCTGCTGCTGCGGCTTACGCTTGT 180  
QY 181 CTACTAACCTGCTGTAAGCGCGCGCGCGCGGCGGCGATGCGCACTGCGGCGCGAC 240  
DB 181 CTACTAACCTGCTGTAAGCGCGCGCGCGCGGCGGCGATGCGCACTGCGGCGCGAC 240  
QY 241 GCGCGGATGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
DB 241 GCGCGGATGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
QY 301 CCGGAGAGCGCGCTGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 360  
DB 301 CCGGAGAGCGCGCTGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 360  
QY 361 CGACTCCCGCAGAGAGAGTGGAAACAATGAGTATCTTATGCTTGAAGCTTGGCAG 420  
DB 361 CGACTCCCGCAGAGAGAGTGGAAACAATGAGTATCTTATGCTTGAAGCTTGGCAG 420  
QY 421 TCGGCGCTGGGCGCGCTTGGCACTGCTTCTGAGCTTGAAGCAAGGTGAGCAT 480  
DB 421 TCGGCGCTGGGCGCGCTTGGCACTGCTTCTGAGCTTGAAGCAAGGTGAGCAT 480  
QY 481 CCTCATCAACATGCGGATCACTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCT 540  
DB 481 CCTCATCAACATGCGGATCACTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCT 540  
QY 541 GCTTGGGATGAAACATATCGGCTGCTTCTGCTGACATCTGCTGCTGCTGCTGAA 600  
DB 541 GCTTGGGATGAAACATATCGGCTGCTTCTGCTGACATCTGCTGCTGCTGCTGAA 600  
QY 601 GCGATGCGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 660  
DB 601 GCGATGCGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 660  
QY 661 TGAATTAAGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720  
DB 661 TGAATTAAGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720  
QY 721 TGAATTAAGAGCTGAGCTGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCT 780  
DB 721 TGAATTAAGAGCTGAGCTGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCT 780  
QY 781 TGGGATGAGCTGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 840

DB 781 TGGGATGAGCTGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 840  
QY 841 TGGTCTGATGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 900  
DB 841 TGGTCTGATGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 900  
QY 901 AAGAGGAGGAGTGGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 960  
DB 901 AAGAGGAGGAGTGGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 960  
QY 961 TGGGATGAGCTGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1020  
DB 961 TGGGATGAGCTGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1020  
QY 1021 GCGAGCGCGCTGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1080  
DB 1021 GCGAGCGCGCTGAGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1080  
QY 1081 TGGTGAACCGATGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1140  
DB 1081 TGGTGAACCGATGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1140  
QY 1141 CCGCGAGCGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1200  
DB 1141 CCGCGAGCGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1200  
QY 1201 TTTGTCTAAGATAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1260  
DB 1201 TTTGTCTAAGATAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1260  
QY 1261 ACCCTGAGCGCGAGAGCTTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCT 1320  
DB 1261 ACCCTGAGCGCGAGAGCTTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCT 1320  
QY 1321 TGAAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1380  
DB 1321 TGAAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1380  
QY 1381 GCGCGCGAGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1440  
DB 1381 GCGCGCGAGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1440  
QY 1441 TCTTCTGAGAGTGAAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCT 1500  
DB 1441 TCTTCTGAGAGTGAAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCT 1500  
QY 1501 TCTGCGAATTTGAGATGATATTTCAAGCGCGCGCGCGCGCGCGCGCGCT 1560  
DB 1501 TCTGCGAATTTGAGATGATATTTCAAGCGCGCGCGCGCGCGCGCGCGCT 1560  
QY 1561 GAGCAGAGCGAGGAGTTCAGATGATGATGATGATGATGATGATGATGATGAT 1620  
DB 1561 GAGCAGAGCGAGGAGTTCAGATGATGATGATGATGATGATGATGATGATGAT 1620  
QY 1621 TCCCTTTCAGACCGCTGAGTATGATGATGATGATGATGATGATGATGATGAT 1680  
DB 1621 TCCCTTTCAGACCGCTGAGTATGATGATGATGATGATGATGATGATGATGAT 1680  
QY 1681 GGTCTCCGAGCTGAGAGCTGCGCTGAGGCTTCAAGTGAAGTGAAGGAGGCT 1740  
DB 1681 GGTCTCCGAGCTGAGAGCTGCGCTGAGGCTTCAAGTGAAGTGAAGGAGGCT 1740  
QY 1741 CTGAACCGAAGGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1800  
DB 1741 CTGAACCGAAGGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1800  
QY 1801 CAGGCGAGGCGAGCTGATGAGGCTGCGCGCGCGCGCGCGCGCGCGCGCT 1860  
DB 1801 CAGGCGAGGCGAGCTGATGAGGCTGCGCGCGCGCGCGCGCGCGCGCGCT 1860  
QY 1861 ATGAGGCGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1920

DB 1861 ATGCGAGGCTGAGGCTGATCTGATGAGAGCCCTCTGAGTAAGAGCGCTTGACCGCCA 1920  
QY 1921 AAAAAAAAAAAAAAAAAA 1939  
DB 1921 AAAAAAAAAAAAAAAAAA 1939

RESULT 3  
AACT8501  
ID AACT8501 standard; cDNA, 1939 BP.  
XX  
AC AACT8501;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Human PRO853 (UNQ419) nucleotide sequence SEQ ID NO:205.  
XX  
KW Human, secreted protein; transmembrane protein; PRO; EST; cytosolic;  
XX expressed sequence tag; detection; cancer; se.  
XX Homo sapiens.  
XX OS  
XX PN WO200053756-A2.  
XX PD 14-SEP-2000.  
XX  
PF 18-FEB-2000; 2000WO-US04341.  
XX  
PR 08-MAR-1999; 99WO-US05028.  
PR 12-MAR-1999; 99US-0123857.  
PR 29-MAR-1999; 99US-0126773.  
PR 21-APR-1999; 99US-0130232.  
PR 28-APR-1999; 99US-0131445.  
PR 14-MAY-1999; 99US-0134287.  
PR 23-JUN-1999; 99US-0141037.  
PR 26-JUL-1999; 99US-0145698.  
PR 29-OCT-1999; 99US-0162506.  
PR 30-NOV-1999; 99WO-US28513.  
PR 02-DEC-1999; 99WO-US28551.  
PR 02-DEC-1999; 99WO-US28565.  
PR 16-DEC-1999; 99WO-US30095.  
PR 30-DEC-1999; 99WO-US31243.  
PR 30-DEC-1999; 99WO-US31274.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
XX  
XX  
XX (GENTH) GENENTECH INC.  
XX  
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL,  
PI Ferrara N, Flivaroff E, Fong S, Gao W, Garber H, Gerritsen ME,  
PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ,  
PI Klijavin IJ, Kuo SS, Napier MA, Pan J, Paoni NP, Roy MA,  
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;  
XX  
XX WPI, 2000-611443/58.  
DR P-PSDB; AAB44271.  
XX  
XX Novel PRO polypeptides and polynucleotides used in detection methods,  
PT to target bioactive molecules to specific cells, and to modulate  
PT cellular activities -  
XX  
XX Claim 2; Fig 74; 636pp; English.  
XX  
XX AACT8458 to AACT8599 represent polynucleotide and EST (expressed  
CC sequence tag) sequences which encode secreted or transmembrane PRO  
CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic  
CC activity. The polynucleotides and polypeptides can be used for detecting  
CC the presence of PRO polypeptides in samples, for linking bioactive  
CC molecules to cells and for modulating biological activities of cells,  
CC using the polypeptides for specific targeting. The polypeptide targeting  
CC can be used to kill the target cells, e.g. for the treatment of cancers.  
CC The polypeptide pairs provide specific targeting of bioactive molecules

CC to cells. AACT8600 to AACT8987 represent PCR primers and probes used in  
CC the isolation of the PRO polynucleotide sequences.  
XX  
SQ Sequence 1939 BP; 367 A; 549 C; 624 G; 399 T; 0 other;  
Query Match 100.0%; Score 1939; DB 21; Length 1939;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1939; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCTCCGCTTGGAGGCTGAGCGCCCGGCGCGCTTCCAGAGCGCTTGACAGGGCGGAT 60  
DB 1 CGCCTCCGCTTGGAGGCTGAGCGCCCGGCGCGCTTCCAGAGCGCTTGACAGGGCGGAT 60  
QY 61 CGGAGCGCGCTTGGAGGCTGAGCGCGGATCAGAGCGGCTGAGCGCGGAGCGCGGAGCGCG 120  
DB 61 CGGAGCGCGCTTGGAGGCTGAGCGCGGATCAGAGCGGCTGAGCGCGGAGCGCGGAGCGCG 120  
QY 121 GCGCGGATGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
DB 121 GCGCGGATGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
QY 181 CTACTACAACTGGAGAGGCG 240  
DB 181 CTACTACAACTGGAGAGGCG 240  
QY 241 GCGCGTGTGTCAGCG 300  
DB 241 GCGCGTGTGTCAGCG 300  
QY 301 CCGGAGAGCG 360  
DB 301 CCGGAGAGCG 360  
QY 361 CGACCTCGCGCGAGAGAGGAG 420  
DB 361 CGACCTCGCGCGAGAGAGGAG 420  
QY 421 TCTGCGCTCGGTCG 480  
DB 421 TCTGCGCTCGGTCG 480  
QY 481 CCTATCCACAAATGCGCGGATATGATTCCTGTCGCGCGCGCGCGCGCGCGCGCGCGCG 540  
DB 481 CCTATCCACAAATGCGCGGATATGATTCCTGTCGCGCGCGCGCGCGCGCGCGCGCGCG 540  
QY 541 GCTTGGGAGAGCAATGCGGATGCGGCTTCTGTCGAGACATGCTGCTGCTGCTGCTGCTG 600  
DB 541 GCTTGGGAGAGCAATGCGGATGCGGCTTCTGTCGAGACATGCTGCTGCTGCTGCTGCTG 600  
QY 601 GGCATGTGCGCCCTAGCGCGGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660  
DB 601 GGCATGTGCGCCCTAGCGCGGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660  
QY 661 TGACTTCAAGCGCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720  
DB 661 TGACTTCAAGCGCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720  
QY 721 TGACACTAGCGGCTGATATGATGCTGTTTCCCGGAGAGCTCGGCGCAACCACTTGAAGCC 780  
DB 721 TGACACTAGCGGCTGATATGATGCTGTTTCCCGGAGAGCTCGGCGCAACCACTTGAAGCC 780  
QY 781 TGGCGTCACTGTGATGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840  
DB 781 TGGCGTCACTGTGATGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840  
QY 841 TGTTCCTGAGAGCTGCG 900  
DB 841 TGTTCCTGAGAGCTGCG 900  
QY 901 AAGAGGGGGTGGCGAGAGCG 960  
DB 901 AAGAGGGGGTGGCGAGAGCG 960

```

QY 961 TGGGAAATATTTTCCCACTGCAATGTGAAAGAGTGCCTTCCAGCTGCCCGAGACGACCG 1020
DB 961 TGGGAAATATTTTCCCACTGCAATGTGAAAGAGTGCCTTCCAGCTGCCCGAGACGACCG 1020
QY 1021 GGGAGCCCATCGGCTATGGGAGGCGCAGCAAGAGGCTGAGTAGGGCTTGGGCGGAGAG 1080
DB 1021 GGGAGCCCATCGGCTATGGGAGGCGCAGCAAGAGGCTGAGTAGGGCTTGGGCGGAGAG 1080
QY 1081 TGGTGAACCCGATGAAGACCCCGAGCTGTGAGACTCAGAGGCCCATCTTCTTAAGCAG 1140
DB 1081 TGGTGAACCCGATGAAGACCCCGAGCTGTGAGACTCAGAGGCCCATCTTCTTAAGCAG 1140
QY 1141 CCCCCACCTGAGAGAGCCCACTGTTCTCACTTACCCGAGCCCTCAGAGCTCAGCAG 1200
DB 1141 CCCCCACCTGAGAGAGCCCACTGTTCTCACTTACCCGAGCCCTCAGAGCTCAGCAG 1200
QY 1201 TTTGTCTAAGATAGCAGCAGGATTCAGAGCTAAGTTGAGGCTGAGATCCAGCTCCCTA 1260
DB 1201 TTTGTCTAAGATAGCAGCAGGATTCAGAGCTAAGTTGAGGCTGAGATCCAGCTCCCTA 1260
QY 1261 ACCCTCAGGCGCAGATGCTGTCATGAGCACTTCTCATGCTCTTGAAGAACTCGAGTGTG 1320
DB 1261 ACCCTCAGGCGCAGATGCTGTCATGAGCACTTCTCATGCTCTTGAAGAACTCGAGTGTG 1320
QY 1321 TGAAGGCAATGCTCTGAGCACTGACGGGTTTGTGATCTTGAACCTCGGTGTTACTTTCTG 1380
DB 1321 TGAAGGCAATGCTCTGAGCACTGACGGGTTTGTGATCTTGAACCTCGGTGTTACTTTCTG 1380
QY 1381 GGGCCCAAGCTGTGCTGAGCACTCTTTCTGCTGTTGAAGGATTAAGGCTGATTT 1440
DB 1381 GGGCCCAAGCTGTGCTGAGCACTCTTTCTGCTGTTGAAGGATTAAGGCTGATTT 1440
QY 1441 TCTTCTGAGATGAGCACTAACCAGATGAGAGATAGAGGATGCTAGACATGCTGCT 1500
DB 1441 TCTTCTGAGATGAGCACTAACCAGATGAGAGATAGAGGATGCTAGACATGCTGCT 1500
QY 1501 TCTTCTGAGATGAGCACTAACCAGATGAGAGATAGAGGATGCTAGACATGCTGCT 1560
DB 1501 TCTTCTGAGATGAGCACTAACCAGATGAGAGATAGAGGATGCTAGACATGCTGCT 1560
QY 1561 GAGCAGAGGCGAGGAGTTGCAATGTGATGAGCACTGCGCAATGAGAAATTAAGTGA 1620
DB 1561 GAGCAGAGGCGAGGAGTTGCAATGTGATGAGCACTGCGCAATGAGAAATTAAGTGA 1620
QY 1621 TCCCTTTGCAACCGCTAGCTAGTATGATTAATTAATTAATTAATTAATTAATTAATTA 1680
DB 1621 TCCCTTTGCAACCGCTAGCTAGTATGATTAATTAATTAATTAATTAATTAATTAATTA 1680
QY 1681 GGGTCCCGAGCTAAGGAGCTCGGCTTGGGCTTCACTGAGTGAAGAGAGGCGCTGGAGT 1740
DB 1681 GGGTCCCGAGCTAAGGAGCTCGGCTTGGGCTTCACTGAGTGAAGAGAGGCGCTGGAGT 1740
QY 1741 CTGAACCCCAAGGCTGAGGCGCAGGCGCAGCTGCTAAGTGAAGTGGGCTGAGAGAGT 1800
DB 1741 CTGAACCCCAAGGCTGAGGCGCAGGCGCAGCTGCTAAGTGAAGTGGGCTGAGAGAGT 1800
QY 1801 CAGGCGAGGCGAGCTGATTCAGAGTGCCTCATGAGAGTGAAGGAGCGCTTCCGCGCG 1860
DB 1801 CAGGCGAGGCGAGCTGATTCAGAGTGCCTCATGAGAGTGAAGGAGCGCTTCCGCGCG 1860
QY 1861 ATGCAAGGCTGAGGCTGATTCAGAGTGCCTCATGAGAGTGAAGGAGCGCTTCCGCGCG 1920
DB 1861 ATGCAAGGCTGAGGCTGATTCAGAGTGCCTCATGAGAGTGAAGGAGCGCTTCCGCGCG 1920
QY 1921 AAAAAAAAAAAAAAAAAA 1939
DB 1921 AAAAAAAAAAAAAAAAAA 1939

```

RESULT 4  
 AAA46932  
 ID AAA46932 standard; cDNA, 1939 BP.  
 XX

```

AC AAA46932;
XX
XX 03-OCT-2000 (first entry)
XX
XX CDNA encoding novel polypeptide PRO853.
DE
XX PRO201, PRO292, PRO327, PRO1265, PRO344, PRO343, PRO347, PRO357,
XX PRO715, PRO1017, PRO1112, PRO509, PRO853, PRO882, tumour cell,
XX tumorigenesis; cancer; neoplastic cell growth; cell proliferation; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 128..1261
XX FT /*tag= a
XX
XX MO200037640-A2.
XX
XX 29-JUN-2000.
XX
XX 16-DEC-1999; 99WO-US30095.
XX
XX 22-DEC-1998; 98US-0113296.
XX 08-MAR-1999; 99WO-US05028.
XX 02-JUN-1999; 99WO-US12252.
XX 01-SEP-1999; 99WO-US20111.
XX 15-SEP-1999; 99WO-US21090.
XX 30-NOV-1999; 99WO-US28313.
XX 01-DEC-1999; 99WO-US28409.
XX 02-DEC-1999; 99WO-US28301.
XX
XX (GETH ) GENENTECH INC.
XX
XX Botstein D, Goddard A, Gurney AL, Hillan K, Lawrence DA, Roy MA,
XX Wood WI;
XX WPI; 2000-452188/39.
XX P-PSDB; AAY93696.
XX
XX New anti-polypeptide antibody useful in the treatment and diagnosis of
XX neoplastic cell growth and proliferation -
XX
XX Claim 50; Fig 25; 220bp; English.
XX
XX The present sequence encodes a novel human polypeptide. The
XX specification describes novel polypeptides designated PRO201, PRO292,
XX PRO327, PRO1265, PRO344, PRO343, PRO347, PRO357, PRO1017,
XX PRO1112, PRO509, PRO853 and PRO882. These genes are amplified in
XX the genome of tumour cells. The polypeptides are believed to contribute
XX to tumorigenesis. The polypeptides are useful target for the
XX identification of certain cancers, and may act as predictors of the
XX prognosis of tumour treatment. Antibodies against these polypeptides
XX are useful in the treatment and diagnosis of neoplastic cell growth
XX and proliferation in mammals.
XX
XX Sequence 1939 BP; 367 A; 549 C; 624 G; 399 T; 0 other;
SQ
XX
XX Query Match 100.0%; Score 1939; DB 21; Length 1939;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1939; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCCTCCGCTTCGAGAGCTGAGCGGCGGCGGCGGCTTCCAGAGCTGTGAGGCGGAGT 60
DB 1 CGCCTCCGCTTCGAGAGCTGAGCGGCGGCGGCGGCTTCCAGAGCTGTGAGGCGGAGT 60
QY 61 CGGCAAGCGGCTGAGCGGCGATTCAGAGCGGCTGAGCGGCGGCGGAGCGGCGGCGG 120
DB 61 CGGCAAGCGGCTGAGCGGCGATTCAGAGCGGCTGAGCGGCGGCGGAGCGGCGGCGG 120
QY 121 GGGCGGCAATGAGAGGCGCTGCTGAGGCGGCGGCGGCTTCTGCTGAGGCGCTTAAGTGTGT 180
DB 121 GGGCGGCAATGAGAGGCGCTGCTGAGGCGGCGGCGGCTTCTGCTGAGGCGCTTAAGTGTGT 180

```

[illegible]

Oy		1261	ACCTCGAGGCGCAGAATCTTCCATGGCACTTAATGTCTCTGAAAACCTCGAATGCTG	1320
Dd		1261	ACCCTTGAGGCCAGGATCTTGCCTTGAAGCACTTAAATGTTACTGAAACCTCGAATGCTG	1320
Oy		1321	TGAGGCCATGCGCTTGACA CTGACGGGTTTTGTGATCTTGA CTTCCGCTGTTACTTTTCGG	1380
Dd		1321	TGAGGCCATGCGCTTGACA CTGACGGGTTTTGTGATCTTGA CTTCCGCTGTTACTTTTCGG	1380
Oy		1381	GCCCCCAAGCTGTGATGCTTGACATCTTTTCTGTGTTGAAGAAATAATGGGTGATTATTT	1440
Dd		1381	GCCCCCAAGCTGTGATGCTTGACATCTTTTCTGTGTTGAAGAAATAATGGGTGATTATTT	1440
Oy		1441	TCTTCTGAGAGTAGACGTAAACCCCAATGAGAGATAGGGGTATGTGTAGACATGTGCT	1500
Dd		1441	TCTTCTGAGAGTAGACGTAAACCCCAATGAGAGATAGGGGTATGTGTAGACATGTGCT	1500
Oy		1501	TCTCGGAAATTTGGATGTAGTATATTTTCAGGCCCCCAACCTTAATTGATCTGATCAGTCTG	1560
Dd		1501	TCTCGGAAATTTGGATGTAGTATATTTTCAGGCCCCCAACCTTAATTGATCTGATCAGTCTG	1560
Oy		1561	GAGCAGAGGCGAGGAGTTTGCAATGTGATSCA CTGCCAACATTGAGAAATTAGTAACTGA	1620
Dd		1561	GAGCAGAGGCGAGGAGTTTGCAATGTGATSCA CTGCCAACATTGAGAAATTAGTAACTGA	1620
Oy		1621	TTCCTTTTGCAACCGTCTAGCTAGTAA GTTAAATTAATCCCCTCATGTATATGAAGCGGAATTA	1680
Dd		1621	TTCCTTTTGCAACCGTCTAGCTAGTAA GTTAAATTAATCCCCTCATGTATATGAAGCGGAATTA	1680
Oy		1681	GGCTCCCGAGGTAA GGGACTCGCCTAGGAGTCTCA CACTGATGATGAGAGAGGGCCTGGGAT	1740
Dd		1681	GGCTCCCGAGGTAA GGGACTCGCCTAGGAGTCTCA CACTGATGATGAGAGAGGGCCTGGGAT	1740
Oy		1741	CTGAACCCCAAAGGCTCTGAGGCGCAGGCGCA CTGCGCTAAGATGGGTGCTGAGAA GTAGT	1800
Dd		1741	CTGAACCCCAAAGGCTCTGAGGCGCAGGCGCA CTGCGCTAAGATGGGTGCTGAGAA GTAGT	1800
Oy		1801	CAGGGCACGGGCACTGTGTATCGAGTCCCA CATGGGAGTAAAGGGAGACGCTTCGGGGCGG	1860
Dd		1801	CAGGGCACGGGCACTGTGTATCGAGTCCCA CATGGGAGTAAAGGGAGACGCTTCGGGGCGG	1860
Oy		1861	ATGCAAGGGGCTGGGGTCA TCTGTATCTGAAGCCCTCGGA ATPAAGCGCGTTGACCGGCCAA	1920
Dd		1861	ATGCAAGGGGCTGGGGTCA TCTGTATCTGAAGCCCTCGGA ATPAAGCGCGTTGACCGGCCAA	1920
Oy		1921	AAAAAAAAAAAAAAAAAAAA 1939	
Dd		1921	AAAAAAAAAAAAAAAAAAAA 1939	

RESULT 5  
 ABX92437 standard; cDNA, 1939 BP.

ABX92437;  
 08-MAY-2003 (first entry)

cDNA encoding human PRO853 polypeptide.

Human, PRO polypeptide; secreted and transmembrane protein; immune disorder; diabetes; hyper-insulinaemia; hypo-insulinaemia; cardiac insufficiency; nervous system disorder; kidney disorder; bone disorder; cartilage disorder; arthritis; tumour; wound healing; genetic disorder; cytostatic; antidiabetic; antiinflammatory; arthralgic; anti-tumour; vulnery; antianemic; dermatological; cardiac; gene; ss.

Homo sapiens.  
 US2002169284-A1.  
 14-NOV-2002.

```
PF 16-OCT-2001; 2001US-0978697.
XX 07-OCT-1998; 98WO-US21141.
PR 20-NOV-1998; 98WO-US24855.
PR 05-JAN-1999; 99WO-US00106.
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99WO-US05190.
PR 14-MAY-1999; 99WO-US05190.
PR 02-JUN-1999; 99WO-US12252.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 16-DEC-1999; 99WO-US30095.
PR 30-DEC-1999; 99WO-US31243.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 22-MAR-2001; 2001WO-US09552.
PR 25-MAY-2001; 2001WO-US17092.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 09-JUL-2001; 2001WO-US21066.
PR 29-JUL-2001; 2001WO-US21735.
PR 17-OCT-1997; 97US-062250P.
PR 03-NOV-1997; 97US-064249P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066364P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077641P.
PR 12-MAR-1998; 98US-077649P.
PR 13-MAR-1998; 98US-077791P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078910P.
PR 20-MAR-1998; 98US-078936P.
PR 20-MAR-1998; 98US-078939P.
PR 25-MAR-1998; 98US-079294P.
PR 26-MAR-1998; 98US-079656P.
PR 27-MAR-1998; 98US-079663P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079683P.
PR 27-MAR-1998; 98US-079728P.
PR 27-MAR-1998; 98US-079786P.
PR 30-MAR-1998; 98US-079920P.
PR 30-MAR-1998; 98US-079923P.
PR 26-MAY-1991; 81US-0267213.
PR 17-MAR-1998; 98US-0040220.
PR 26-JUN-1998; 98US-0105413.
PR 07-OCT-1998; 98US-0168978.
PR 02-NOV-1998; 98US-0184216.
PR 06-NOV-1998; 98US-0187368.
PR 07-DEC-1998; 98US-0202054.
PR 22-DEC-1998; 98US-0218517.
PR 05-MAR-1999; 99US-0254465.
PR 10-MAR-1999; 99US-0265686.

PR 12-APR-1999; 99US-0284291.
PR 14-MAY-1999; 99US-0311832.
PR 14-MAY-1999; 99US-0380137.
PR 25-AUG-1999; 99US-0380138.
PR 25-AUG-1999; 99US-0380142.
PR 08-NOV-2000; 2000US-0709238.
PR 27-NOV-2000; 2000US-0723749.
PR 20-DEC-2000; 2000US-0747259.
PR 22-MAR-2001; 2001US-0816744.
PR 10-MAY-2001; 2001US-0816920.
PR 10-MAY-2001; 2001US-0854208.
PR 01-JUN-2001; 2001US-0854280.
PR 05-JUN-2001; 2001US-0872035.
PR 14-JUN-2001; 2001US-0882636.
PR 19-JUN-2001; 2001US-0886342.
PR 30-JUL-2001; 2001US-0918585.

(PGTH ) GENENTECH INC.
PI Ashkenazi A, Baker KP, Botstein D, Desnoyers L, Eaton D;
PI Perrera N, Flyvareff E, Fong S, Gao W, Garber H, Gertlisen ME;
PI Goddard A, Goddard PJ, Grimaldi JC, Gunney AL, Hillan KJ;
PI Kijavyn IJ, Kuo SS, Napier MA, Pen J, Paoi NF, Roy MA;
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;
XX WPI; 2003-288163/28.
DR P-PSDB; ABU61101.
XX
PT Novel secreted and transmembrane polypeptides and polynucleotides
PT encoding them useful for treating cancer, kidney diseases, bone,
PT cartilage disorders and immune deficiencies
XX
PS Claim 2; Fig 74; 45pp; English.
XX
CC The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The
CC PRO polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides are useful for detecting other PRO polypeptides, for
CC linking bioactive molecules to cells expressing PRO polypeptides,
CC for modulating biological activities of cells expressing PRO
CC polypeptides, and for identifying agonists or antagonists. The
CC bioactive molecule may be a toxin, radiolabel or antibody, and causes
CC apoptosis or death of the cell. The PRO polypeptides are useful for
CC treating immune disorders, diabetes or hyper- or hypo-insulinaemia,
CC cardiac insufficiency, nervous system disorders, kidney disorders,
CC bone and cartilage disorders or arthritis, tumours, and wound healing.
CC The polynucleotide sequences encoding PRO polypeptides are useful as
CC hybridisation probes, in chromosome and gene mapping, in the generation
CC of antisense RNA and DNA, in the preparation of PRO polypeptides, for
CC generating transgenic animals or knockout animals, for the genetic
CC analysis of individuals with genetic disorders, and in gene therapy.
CC The present sequence encodes a human PRO polypeptide of the invention.
CC Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at
CC seqdata.uspto.gov/patident.html.
CC
XX
SQ Sequence 1939 BP; 367 A; 549 C; 624 G; 399 T; 0 other;

Query Match 100.0%; Score 1939; DB 25; Length 1939;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1939; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCTCGCCTTGGAGAGCTGACGCGCGCGCGCGCTTCCAGGCTGTGACAGGCGGAT 60
Db 1 CGCCTCGCCTTGGAGAGCTGACGCGCGCGCGCGCTTCCAGGCTGTGACAGGCGGAT 60
QY 61 CGGACGCGCTGCGCGCGGATCCAGGCGCGGCTGTGCGGCGCTGTGCGGAGCGCGGCGC 120
Db 61 CGGACGCGCTGCGCGCGGATCCAGGCGCGGCTGTGCGGCGCTGTGCGGAGCGCGGCGC 120
QY 121 GGGCGGATGAGAGCGCGCTGTGCTGGCGGCGGCGGCTGTGCGGCGCTTACGCTTGT 180
Db 121 GGGCGGATGAGAGCGCGCTGTGCTGGCGGCGGCGGCTGTGCGGCGCTTACGCTTGT 180
```





XX 19-MAY-2000; 2000US-205515P.  
 2R

2A (HUMA-) HUMAN GENOME SCI INC.

PI Blrse CE, Rosen CA,  
4Y

DR WPI; 2002-122018/16.  
DR P-24DA: ABB89660

XX  
PT  
Novel 1405 isolat

PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
PT prevention of neural, immune system, muscular, reproductive and  
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
PT disorders -  
XX  
XX  
XX Claim 4; SEQ ID NO 631; 2081pp + Sequence Listing; English.  
PS

Claim 4; SEQ ID NO 631; 2081pp + Sequence Listing; English.

CC The invention relates to novel genes (AB089449-AB090553) and proteins  
CC (AB089040-AB089444) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification...  
CC The nucleic acids, proteins, antibodies and (anti)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at [ftp.wipo.int/publ/published\\_pat\\_sequences](http://ftp.wipo.int/publ/published_pat_sequences).  
XX  
XX Sequence 1887 BP; 378 A; 522 C; 597 G; 390 T; 0 other.

**SQ** Sequence 1887 BP; 378 A; 522 C; 597 G; 390 T; 0 other;

Query Match	93.0%	Score 1803	DB 24	Length 1867
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1953	Conservative	0	Mismatches 1	Indels 0
			Gaps	0

QY	86	GCGCGTGGGGGACCTGGGCGGGGAGCGGGAGGCGGGCCGCGATGAGAGGCGCTGGTGGCTG	145
Db	14	GCGCGTGGGGGACCTGGGCGGGGAGCGGGAGGCGGGCCGCGATGAGAGGCGCTGGTGGCTG	73
QY	146	GCGCGGGGGTGTCTGTGGGCGCTTACGTGCTTGTCTACCTACCACTGTGTGAAGCCCCG	205
Db	74	GCGCGGGGGTGTCTGTGGGCGCTTACGTGCTTGTCTACCACTGTGTGAAGCCCCG	133
QY	206	CCGTGCGCGGCGCATGGGGGACCTTGGGGGGCCGACACCGCCCTGTGTCAACGGGCCCAACAC	265
Db	134	CCGTGCGCGGCGCATGGGGGACCTTGGGGGGCCGACACCGCCCTGTGTCAACGGGCCCAACAC	193
QY	266	GGCATCGGAAGAATGACCGCGCTGTGAGCTGCGCGCCCGGGAGACGCGCGTGTGTGCTGGCC	325
Db	194	GGCATCGGAAGAATGACCGCGCTGTGAGCTGCGCGCCCGGGAGACGCGCGTGTGTGCTGGCC	253
QY	326	TGCGCGACGCCAGAGACGCGGGGAGCGCGGCTGCTTCGACCTTCGCCAGAGAGTGGGAAC	385
Db	254	TGCGCGACGCCAGAGACGCGGGGAGCGCGGCTGCTTCGACCTTCGCCAGAGAGTGGGAAC	313
QY	386	AATGAGGTCATCTTCATGAGCCCTTGGACCTTGGCAGTCGTGAGCCCTGGGTGGCGGCCCTTTGGC	445
Db	314	AATGAGGTCATCTTCATGAGCCCTTGGACCTTGGCAGTCGTGAGCCCTGGGTGGCGGCCCTTTGGC	373
QY	446	ACTGCGCTTTCGAGCTCTGAGGCCAGCGTGTGACATCTCTATTCACAACAATGCGCGTATCACT	505
Db	374	ACTGCGCTTTCGAGCTCTGAGGCCAGCGTGTGACATCTCTATTCACAACAATGCGCGTATCACT	433
QY	506	TCTCTGTGGCCGGAACCCGTGAGGCGCTTTAACTGTGTGCTTCGAGGTGAACCATATCGGTGCC	565
Db	434	TCTCTGTGGCCGGAACCCGTGAGGCGCTTTAACTGTGTGCTTCGAGGTGAACCATATCGGTGCC	493

[illegible]

[illegible]

diagnosing, treating and/or preventing autoimmune, inflammatory, cell proliferative, developmental, endocrine, eye, metabolic, and gastrointestinal disorders -

Claim 5; Page 160; 133pp; English.

The present sequence is human drug metabolizing enzyme (DME-3) cDNA. Human DME and its nucleic acid molecule are useful for the diagnosis, treatment and prevention of disorders associated with increased or decreased expression of DME. Examples of such disorders include, autoimmune/inflammatory disorder such as acquired immune deficiency syndrome (AIDS), rheumatoid arthritis, osteoporosis; cell proliferative disorder such as actinic keratosis, atherosclerosis; developmental disorder such as epilepsy, anemia; endocrine disorder such as acromegaly, cretinism, thyrotoxicosis; pancreatic disorder such as diabetes mellitus; eye disorder such as conjunctivitis, glaucoma, iritis; metabolic disorder such as Addison's disease, obesity; gastrointestinal disorder such as anorexia, dysphagia and hepatic tumours including nodular hyperplasia, adenomas and carcinomas. DME DNA is useful for creating "knockin" humanised animals (pigs) or transgenic animals (mice or rats) to model human disease. DME DNA is also in useful in gene therapy. DME and its immunogenic fragments are useful for screening libraries of compounds in several drug screening assays.

Sequence 1779 BP; 385 A; 475 C; 523 G; 396 T; 0 other;

Query Match 77.2%; Score 1497; DB 22; Length 1779;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1547; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY	311	CAGAGAGGTGGAAACATGAGTCACTTCAATGAGCCTTGAACCTTGACAGTCTGGCCTCG	430
DB	222	CAGAGAGGTGGAAACATGAGTCACTTCAATGAGCCTTGAACCTTGACAGTCTGGCCTCG	281
QY	431	GTGGGGGGCTTGGCACTGACCTTTCTGAGCTCTGAGACGAGGTTGAGATCCTCATCCAC	490
DB	282	GTGGGGGGCTTGGCACTGACCTTTCTGAGCTCTGAGACGAGGTTGAGATCCTCATCCAC	341
QY	491	AATGGCGGTATCAGTTCCTGTGAGCGGACCCGTGAGGCGTTAACTGTCTCGGGTG	550
DB	342	AATGGCGGTATCAGTTCCTGTGAGCGGACCCGTGAGGCGTTAACTGTCTCGGGTG	401
QY	551	AACCAATATCGGTCCCTTTCTGCTGACACATCTGCTGCTGCTTGCTGAGAGCATGTGCC	610
DB	402	AACCAATATCGGTCCCTTTCTGCTGACACATCTGCTGCTGCTTGCTGAGAGCATGTGCC	461
QY	611	CTTAGCCGCGGT	670
DB	462	CTTAGCCGCGGT	521
QY	671	CGCCTGACGCGCCAGT	730
DB	522	CGCCTGACGCGCCAGT	581
QY	731	CTGGCTATATGATCTGTTTGGCCGGAGCTCCCAACAGCTTGAGGCGCACTGCGTCAAC	790
DB	582	CTGGCTATATGATCTGTTTGGCCGGAGCTCCCAACAGCTTGAGGCGCACTGCGTCAAC	641
QY	791	TGCTATGACAGCCCAACCAAGGCGCTGTGTAACTGAGAGCTGTTCTGCGCATGTTCTGGA	850
DB	642	TGCTATGACAGCCCAACCAAGGCGCTGTGTAACTGAGAGCTGTTCTGCGCATGTTCTGGA	701
QY	851	TGGCGTGGCGCACTTTTGGCGCCCAATGTGGCTTGGTGTGCTCGGGGCAACAAAGGGGGT	910
DB	702	TGGCGTGGCGCACTTTTGGCGCCCAATGTGGCTTGGTGTGCTCGGGGCAACAAAGGGGGT	761
QY	911	GCCAGACACCCCTGTATTTGTCTCTACAGAGGCGCATCGAGCCCTCAGTGGAGATAT	970
DB	762	GCCAGACACCCCTGTATTTGTCTCTACAGAGGCGCATCGAGCCCTCAGTGGAGATAT	821
QY	971	TTTGGCAACTGCTGATGTGAAAGAGGTGCTTCAGGTGCTCGAGACAGACCGGGCAGCCCAT	1030
DB	822	TTTGGCAACTGCTGATGTGAAAGAGGTGCTTCAGGTGCTCGAGACAGACCGGGCAGCCCAT	881

QY 1031 CGGCTATGGAGGCCAGCAAGAGCTGGCAAGGCTTGGGCTGGGGAGAGTCTGAACC 1090  
DB 882 CGGCTATGGAGGCCAGCAAGAGCTGGCAAGGCTTGGGCTGGGGAGAGTCTGAACC 941  
QY 1091 GATGAGAGCCCGGAGCTGAGAGTCAAGAGGCGCCCATCTTCTTAAGCAACCCCACTT 1150  
DB 942 GATGAGAGCCCGGAGCTGAGAGTCAAGAGGCGCCCATCTTCTTAAGCAACCCCACTT 1001  
QY 1151 GAGAGGCCCGGAGCTGAGAGTCAAGAGGCGCCCATCTTCTTAAGCAACCCCACTT 1210  
DB 1002 GAGAGGCCCGGAGCTGAGAGTCAAGAGGCGCCCATCTTCTTAAGCAACCCCACTT 1061  
QY 1211 ATGAGCGCACGAAATTGAGGCTAAAGTTGAGGCTGAGAGTCAAGAGTCTTCTTAAGC 1270  
DB 1062 ATGAGCGCACGAAATTGAGGCTAAAGTTGAGGCTGAGAGTCAAGAGTCTTCTTAAGC 1121  
QY 1271 CAGGATGCTTGGCAGTGGCACTTCAAGTCTTGAAGAACTTGGATGTTGTGAGGCTAG 1330  
DB 1122 CAGGATGCTTGGCAGTGGCACTTCAAGTCTTGAAGAACTTGGATGTTGTGAGGCTAG 1181  
QY 1331 CCCTGGACACTGACGCGGTTTGGATCTTGAACCTCCGAGTTACTTTCTGGGGCCCAAGC 1390  
DB 1182 CCCTGGACACTGACGCGGTTTGGATCTTGAACCTCCGAGTTACTTTCTGGGGCCCAAGC 1241  
QY 1391 TGGTCCCTGGACACTGCTCTTCTTCTGTTGAAGAAATATGGGTATTTCTTCTGAG 1450  
DB 1242 TGGTCCCTGGACACTGCTCTTCTTCTGTTGAAGAAATATGGGTATTTCTTCTGAG 1301  
QY 1451 AGGACAGTAAACCCGAGTGGAGAGTAAAGGAGTAAAGTAAAGTAAAGTAAAGTAAAG 1510  
DB 1302 AGGACAGTAAACCCGAGTGGAGAGTAAAGGAGTAAAGTAAAGTAAAGTAAAGTAAAG 1361  
QY 1511 TTGGATGATGATTTTCAAGGCCCACTTATTTGATTTGATTTGATTTGATTTGATTTGAT 1570  
DB 1362 TTGGATGATGATTTTCAAGGCCCACTTATTTGATTTGATTTGATTTGATTTGATTTGAT 1421  
QY 1571 AGGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1630  
DB 1422 AGGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1481  
QY 1631 ACCGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1690  
DB 1482 ACCGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1541  
QY 1691 CTAAAGGACTGCGCTGAGGCTGCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1750  
DB 1542 CTAAAGGACTGCGCTGAGGCTGCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1601  
QY 1751 GGGTCTGAGGCCAGGCCCACTGCTGCTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1810  
DB 1602 GGGTCTGAGGCCAGGCCCACTGCTGCTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1661  
QY 1811 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1870  
DB 1662 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1721  
QY 1871 GGGGTATCTGATCTGAAGCCCTCGGATTAAGGCTTTGACGCC 1918  
DB 1722 GGGGTATCTGATCTGAAGCCCTCGGATTAAGGCTTTGACGCC 1769

RESULT 8  
ABZ68100  
ID ABZ68100 standard, DNA, 1098 BP.  
XX  
AC ABZ68100;  
XX  
XX 26-MAR-2003 (first entry)  
XX Human secreted protein encoding genomic DNA SEQ ID NO 1623.  
XX Human; secreted protein; nootropic; neuroprotective; cytostatic;  
KM

KM virucide; dermatological; immunosuppressive; anti-inflammatory; anti-HIV;  
KM vulnery; antibacterial; antiparkinsonian; antisticking; antianemic;  
KM antihistaminic; cancer; antirheumatic; hepatotropic; cerebroprotective;  
KM antineoplastic; antiallergic; antidiabetic; anticancer; anticonvulsant;  
KM antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;  
KM cardiovascular disorder; neurological disease; neurotropic;  
KM gene therapy; gene; ds.  
OS Homo sapiens.  
XX  
XX MO200277166-A2.  
PN  
XX  
XX 03-OCT-2002.  
PD  
XX  
XX 26-MAR-2002; 2002WO-US09188.  
PF  
XX  
XX 27-MAR-2001; 2001US-278650P.  
PR 12-SEP-2001; 2001US-0950082.  
PR 12-SEP-2001; 2001US-0950083.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX Rosen CA, Ruben SM,  
XX WPI; 2003-040583/03.  
DR  
XX  
XX New human secreted proteins encoded by genes contained in cDNA clones  
PT (e.g. HGC19), useful for preventing, treating or diagnosing e.g.  
PT AIDS, multiple sclerosis, herpes virus, leukemia, tick-borne  
PT encephalitis or West Nile fever.  
XX  
XX Disclosure; Page 2253-2254; 2423pp; English.  
XX  
XX The invention relates to novel human genes (ABZ6891-ABZ68209) and the  
CC encoded secreted proteins (ABP9470-ABP9872) useful for preventing,  
CC treating or ameliorating medical conditions e.g. by protein or gene  
CC therapy. The genes are isolated from a range of human tissues disclosed  
CC in the specification. The nucleic acids, proteins, antibodies and  
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:  
CC (a) cancer; e.g. breast and ovarian cancer and other cancers of the  
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,  
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as  
CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.  
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,  
CC bacterial, fungal and parasitic infections.  
XX  
XX Sequence 1098 BP; 243 A; 286 C; 326 G; 243 T; 0 other;  
SQ

Query Match 40.2%; Score 779; DB 25; Length 1098;  
Best Local Similarity 99.6%; Pred. No. 2.3e-293;  
Matches 979; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 936 TACAAGAGGCGATCGAGCCCTCACTGAGAGTATTTTCCAACTGCGATGAGAGAG 995  
DB 113 TACAAGAGGCGATCGAGCCCTCACTGAGAGTATTTTCCAACTGCGATGAGAGAG 172  
QY 996 TCCCTCCAGCTGCCGAGAGCAGCCGAGCCCATCGGTTATGAGAGGCGAGAGAGC 1055  
DB 173 TCCCTCCAGCTGCCGAGAGCAGCCGAGCCCATCGGTTATGAGAGGCGAGAGAGC 232  
QY 1056 TGGCAGGCGCTTGGGCTGGGAGAGTGTGAACCCGATGAAGAGCCCGAGTGAAGACT 1115  
DB 233 TGGCAGGCGCTTGGGCTGGGAGAGTGTGAACCCGATGAAGAGCCCGAGTGAAGACT 292  
QY 1116 CAGAGGCCCATCTTCTTAAGACACCCCGACCTGAGAGGCCCAAGTTTCAACCTT 1175  
DB 293 CAGAGGCCCATCTTCTTAAGACACCCCGACCTGAGAGGCCCAAGTTTCAACCTT 352  
QY 1176 ACCCGAGCCCTCAAGAGCTCAACAGTTTGTCTTAATGAGCAGCAGCAATTCAGGCTAAG 1235

```

Db 353 ACCGAGCCCTCAGAGCTCACACAGATTGTCTTATAGATACGACCCGATTCAGGCTAAG 412
Cy 1236 TTGAGCTGAGATCCAGCTCTCTTAACCTCCAGGCGCAGAGATGCTTCCAGATGCTTCAAT 1295
Db 443 TTGAGCTGAGATCCAGCTCTCTTAACCTCCAGGCGCAGAGATGCTTCCAGATGCTTCAAT 472
Cy 1296 GGTCTGTAAGAACTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1355
Db 473 GGTCTGTAAGAACTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 532
Cy 1356 CTGAGCTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1415
Db 533 CTGAGCTCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 592
Cy 1416 GTTGAAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1475
Db 593 GTTGAAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 652
Cy 1476 ATAGGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1535
Db 653 ATAGGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 712
Cy 1536 CCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1595
Db 713 CCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 772
Cy 1596 CCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1655
Db 773 CCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 832
Cy 1656 CCCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1715
Db 833 CCCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 892
Cy 1716 AGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1775
Db 893 AGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 952
Cy 1776 GTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1835
Db 953 GTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1012
Cy 1836 GAGTAAGGGGACGCTCCGCGGCGGATGATGATGATGATGATGATGATGATGATGAT 1895
Db 1013 GAGTAAGGGGACGCTCCGCGGCGGATGATGATGATGATGATGATGATGATGATGAT 1072
Cy 1896 CGGAATTAAGCGCTTGAACCGCC 1918
Db 1073 CGGAATTAAGCGCTTGAACCGCC 1095

```

RESULT 9  
AAF26578  
ID AAF26578 standard; DNA; 1140 BP.

AA26578;  
27-MAR-2001 (first entry)  
DNA encoding human secreted protein #32.  
Secreted protein; gene therapy; vaccine; cancer; leukemia;  
autoimmune disease; allergy; inflammation; graft rejection;  
hyperproliferation; cardiovascular; infection; su.  
Homo sapiens.  
WO200076531-A1.  
21-DEC-2000.  
01-JUN-2000; 2000MO-US15137.

```

PR 11-JUN-1999; 99US-0138625.
XX (HUMA-) HUMAN GENOME SCI INC.
XX ROSEN CA, RUBEN SM, KOMATSOUIS GA;
XX WPI; 2001-071148/08.
DR Nucleic acids encoding 47 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -
XX Claim 1, Page 463; 525pp; English.
XX The present invention relates to 26 secreted human proteins. The
CC proteins may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate polypeptide expression.
CC For example, they may be used in gene therapy or in vaccines.
CC Typical of diseases which are potentially treatable are cancers
CC (including leukemia), autoimmune diseases, allergies, inflammation,
CC graft rejection, hyperproliferation, cardiovascular diseases
CC (particularly critical limb ischemia and coronary disease) and any
CC involving abnormal angiogenesis, neurodegeneration and/or
CC infectious diseases.
SQ Sequence 1140 BP; 269 A; 297 C; 330 G; 244 T; 0 other;
Query Match 36.0%; Score 698; DB 22; Length 1140;
Best Local Similarity 99.5%; Pred. No. 7.3e-262;
Matches 1088; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
Cy 827 CTGTTCTCGCCCATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 886
Db 13 CTGTTCTCGCCCATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 72
Cy 887 GTGCTCCGGGACCAAGAGGGGGTCCAGACACCCCTGATGATGATGATGATGATGATGAT 946
Db 73 GTGCTCCGGGACCAAGAGGGGGTCCAGACACCCCTGATGATGATGATGATGATGATGAT 132
Cy 947 ATCGAGCCCTCAGTGGGAGATATTTGCAATGTCAGATGAGAGAGTGCCTCAGCT 1006
Db 133 ATCGAGCCCTCAGTGGGAGATATTTGCAATGTCAGATGAGAGAGTGCCTCAGCT 192
Cy 1007 GCCGAGACGACCGGCGAGCCCATGCGTATGAGAGGCGCAGACAGGCTGCGAGGCTT 1066
Db 193 GCCGAGACGACCGGCGAGCCCATGCGTATGAGAGGCGCAGACAGGCTGCGAGGCTT 252
Cy 1067 GGGCTGGGAGAGATGCTGAACCCGATGAACCCCGATGAGAGATGAGAGGCCCCA 1126
Db 253 GGGCTGGGAGAGATGCTGAACCCGATGAACCCCGATGAGAGATGAGAGGCCCCA 312
Cy 1127 TCTTCTTAAGACACCCCCACCTGAGAGGCCACAGTTTCAACCTTACCCAGCCCT 1186
Db 313 TCTTCTTAAGACACCCCCACCTGAGAGGCCACAGTTTCAACCTTACCCAGCCCT 372
Cy 1187 CAGAGCTCAGCAGATTGTCTTAAGATGACGACCCGATTAAGGCTTAAGGCTGAG 1246
Db 373 CAGAGCTCAGCAGATTGTCTTAAGATGACGACCCGATTAAGGCTTAAGGCTGAG 432
Cy 1247 ATTCAGCTCTCTTAACCTCAGGCGCAGATGCTTGCATGACCTTATGCTTGA 1306
Db 433 ATTCAGCTCTCTTAACCTCAGGCGCAGATGCTTGCATGACCTTATGCTTGA 492
Cy 1307 ACCTCGATGATGATGAGAGCCATGCGCTGACACCTGAGAGGGTTTGTATCTTGA 1366
Db 493 ACCTCGATGATGATGAGAGCCATGCGCTGACACCTGAGAGGGTTTGTATCTTGA 551
Cy 1367 TGGTTACTTTCTGAGG-CCCAAGCTGTGCCCTGACATCTTTCTGCTTGAAGAA 1425
Db 552 TGGTTACTTTCTGAGGCCCCCAAGCTGTGCCCTGACATCTTTCTGCTTGAAGAA 611
Cy 1426 TAAAGGATATTTCTTCTGAGAGTGAACGATGATGATGATGATGATGATGATGATGAT 1485

```

Db 612 TAATGGGATGATTTTCTTCTGAGAGTACACGACCCGACGATGAGAGATAGGGTAT 671  
 QY 1486 GCTAGACACTGTGCTTCTCGAATTTGATGATAGTATTTTCAAGCCCACTATTTGA 1545  
 Db 672 GCTAGACACTGTGCTTCTCGAATTTGATGATAGTATTTTCAAGCCCACTATTTGA 731  
 QY 1546 TTCTGATCAGCTCTGAGACAGAGGAGATTTGCAATGATGACATGCCAATTTGA 1605  
 Db 732 TTCTGATCAGCTCTGAGACAGAGGAGATTTGCAATGATGACATGCCAATTTGA 791  
 QY 1606 GAATAGTGAAGTATGCTTCTTGGCAACGCTAGCTAGTATGATTAATTTACCCCATGTT 1665  
 Db 792 GAATAGTGAAGTATGCTTCTTGGCAACGCTAGCTAGTATGATTAATTTACCCCATGTT 851  
 QY 1666 AATGAGGCGAATTAAGGCTTCCGAGCTAAGGAGACTCGCTAGGCTCTCAAGTATGAG 1725  
 Db 852 AATGAGGCGAATTAAGGCTTCCGAGCTAAGGAGACTCGCTAGGCTCTCAAGTATGAG 911  
 QY 1726 AGGAGGCGCTGGAGATCTGAACCCAGAGGCTCTGAGGCGCAAGGCGCTAAGATGAG 1785  
 Db 912 AGGAGGCGCTGGAGATCTGAACCCAGAGGCTCTGAGGCGCAAGGCGCTAAGATGAG 971  
 QY 1786 TGCTGAGAGTGAAGTACAGGAGGAGGAGCTGATGAGAGTGGCCCATGAGAGTAAAGG 1845  
 Db 972 TGCTGAGAGTGAAGTACAGGAGGAGGAGCTGATGAGAGTGGCCCATGAGAGTAAAGG 1031  
 QY 1846 ACGGCTTCCGAGCGAGATGAGGAGCTGAGGATCTGATCTGATGAGAGGCTTGAATTAAG 1905  
 Db 1032 ACGGCTTCCGAGCGAGATGAGGAGCTGAGGATCTGATCTGATGAGAGGCTTGAATTAAG 1091  
 QY 1906 CGGCTTGAACCGCC 1918  
 Db 1092 CGGCTTGAACCGCC 1104

## RESULT 10

ABZ67229 standard; cDNA, 1140 BP.

ABZ67229;

26-MAR-2003 (first entry)

Human secreted protein encoding cDNA SEQ ID NO 349.

Human; secreted protein; nootropic; neuroprotective; cytosolic;  
 virucide; dermatological; immunosuppressive; anti-inflammatory; anti-HIV;  
 vulnerability; antibacterial; antiparkinsonian; antisticking; antianemic;  
 antiarthritic; cancer; antirheumatic; hepatocytic; cerebroprotective;  
 antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant;  
 antifungal; antiparasitic; cardiac; immune disorder; infection; vaccine;  
 cardiovascular disorder; neurological disease; nephroprotective;  
 gene therapy; gene; ds.

OS Homo sapiens.

PN WO20027186-A2.

03-OCT-2002.

26-MAR-2002; 2002WO-US09188.

27-MAR-2001; 2001US-278650P.

12-SEP-2001; 2001US-0950082.

12-SEP-2001; 2001US-0950083.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI, 2003-040583/03.

P-PsDB; ABP99808.

New human secreted proteins encoded by genes contained in cDNA clones  
 (e.g. HGAC19), useful for preventing, treating or diagnosing e.g.  
 AIDS, multiple sclerosis, herpes virus, leukemia, tick-borne  
 encephalitis or West Nile fever.

Claim 7, Page 1350-1351, 2423pp, English.

The invention relates to novel human genes (ABZ66891-ABZ68209) and the  
 encoded secreted proteins (ABP99470-ABP99872) useful for preventing,  
 treating or ameliorating medical conditions e.g. by protein or gene  
 therapy. The genes are isolated from a range of human tissues disclosed  
 in the specification. The nucleic acids, proteins, antibodies and  
 (ant)agonists are useful in the diagnosis, treatment and prevention of:  
 (a) cancer, e.g. breast and ovarian cancer and other cancers of the  
 adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
 lung or urogenital; (b) immune disorders e.g. Addison's disease,  
 allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
 diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
 arthritis and ulcerative colitis; (c) cardiovascular disorders such as  
 myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.  
 cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,  
 bacterial, fungal and parasitic infections.

Sequence 1140 BP, 269 A, 297 C, 330 G, 244 T, 0 other;

Query Match 36.0%; Score 698; DB 25; Length 1140;

Best Local Similarity 99.5%; Pred. No. 7,36-262; Mismatches 3; Indels 2; Gaps 2;

Matches 1086; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 827 CTGTTCTGCGGCGCCATGTTCTGAGATGCTGGGCCCATTTTGGCCCATTTGGCTTGCTG 886  
 Db 13 CTGTTCTGCGGCGCCATGTTCTGAGATGCTGGGCCCATTTTGGCCCATTTGGCTTGCTG 72  
 QY 887 GTGCTCCGGGCGCCAGAGAGGGGGTGGCCAGACACCCCTGATTTGCTCTCAAGAGAGGCG 946  
 Db 73 GTGCTCCGGGCGCCAGAGAGGGGGTGGCCAGACACCCCTGATTTGCTCTCAAGAGAGGCG 132  
 QY 947 ATCGAGCCCTCAGTGGAGATATTTTGCACCACTGCAATGAGAGAGTGCCTTCAGCT 1006  
 Db 133 ATCGAGCCCTCAGTGGAGATATTTTGCACCACTGCAATGAGAGAGTGCCTTCAGCT 192  
 QY 1007 GCGCGAGACGACCGGGCGAGCCCATGCTATGAGAGGCGACAGAGGCTGGCAGGGCTT 1066  
 Db 193 GCGCGAGACGACCGGGCGAGCCCATGCTATGAGAGGCGACAGAGGCTGGCAGGGCTT 252  
 QY 1067 GGGCTGGGGAGAGTGTGAACCCGATGAAGACCCCGATGAGGATCTCAGAGGCCCA 1126  
 Db 253 GGGCTGGGGAGAGTGTGAACCCGATGAAGACCCCGATGAGGATCTCAGAGGCCCA 312  
 QY 1127 TCTTCTTAAGCAACCCCAACCTGAGAGGCCCAAGTTTCTCAACCTTACCCAGCCCT 1186  
 Db 313 TCTTCTTAAGCAACCCCAACCTGAGAGGCCCAAGTTTCTCAACCTTACCCAGCCCT 372  
 QY 1187 CAGAGCTCAGCAATTTGTCTAAGATGACGACCGAATTCAGGCTTAAGTTGAGCCTGAG 1246  
 Db 373 CAGAGCTCAGCAATTTGTCTAAGATGACGACCGAATTCAGGCTTAAGTTGAGCCTGAG 432  
 QY 1247 ATCGAGCTCTTAACCTGAGGCGAGAGTCTTGCAATGACATCTTATGCTTGA 1306  
 Db 433 ATCGAGCTCTTAACCTGAGGCGAGAGTCTTGCAATGACATCTTATGCTTGA 492  
 QY 1307 ACTCGAGATGTGTGAGGCCATCCCTGAGACATGACGAGGTTTGTGATCTTGA 1366  
 Db 493 ACTCGAGATGTGTGAGGCCATCCCTGAGACATGACGAGGTTTGTGATCTTGA 551  
 QY 1367 TGGTTACTTTCTGGGG-CCCGAAGCTGTGCTTGTGATCTTCTTCTGTTGAAGAA 1425  
 Db 552 TGGTTACTTTCTGGGGCCCGAAGCTGTGCTTGTGATCTTCTTCTGTTGAAGAA 611  
 QY 1426 TAATGGGATTAATTTCTTCTGAGAGTGAAGTAAACCCAGATGAGAGATAGGGGTAT 1485  
 Db 612 TAATGGGATTAATTTCTTCTGAGAGTGAAGTAAACCCAGATGAGAGATAGGGGTAT 671

QY	1486	GCTGAGCACTGTGCTTCGGAAATTGGATGTAGTATTTTCAGGCCACCCCTATTGA	1545
Db	672	GCTGAGCACTGTGCTTCGGAAATTGGATGTAGTATTTTCAGGCCACCCCTATTGA	731
QY	1546	TTCTGATCAGCTTCGAGCAGAGCGAGGATTTGCATGTGATGCATCCGCAATTGA	1609
Db	732	TTCTGATCAGCTTCGAGCAGAGCGAGGATTTGCATGTGATGCATCCGCAATTGA	791
QY	1606	GAATTAGTGAATCTATCCCTTTGCAACCGTCTAGTATGTATTAAATTACCCCATGTT	1665
Db	792	GAATTAGTGAATCTATCCCTTTGCAACCGTCTAGTATGTATTAAATTACCCCATGTT	851
QY	1666	AATGAGCGGATTAAGGCTCCCGAGCTTAAGGAGACTCGCTAGAGGCTCAACAGTAGTAGG	1725
Db	852	AATGAGCGGATTAAGGCTCCCGAGCTTAAGGAGACTCGCTAGAGGCTCAACAGTAGTAGG	911
QY	1726	AGAGAGGCTTCGGATCTGAATCCCAAGGGTCTGAGGCCAGAGGCCGACTGCCCTAAGATGAGG	1785
Db	912	AGAGAGGCTTCGGATCTGAATCCCAAGGGTCTGAGGCCAGAGGCCGACTGCCCTAAGATGAGG	971
QY	1786	TGCTGAGAACTGAGTCAAGGGCAGAGGGCAGCTGGTATCGAGGTGCCCCATAGGGAGTAAGGGG	1845
Db	972	TGCTGAGAACTGAGTCAAGGGCAGAGGGCAGCTGGTATCGAGGTGCCCCATAGGGAGTAAGGGG	1031
QY	1846	ACGCTTTCCGGGCGGATGCAGAGGCTGAGGGTCACTGTATCTGAAGCCCTTCGAGATAAAG	1905
Db	1032	ACGCTTTCCGGGCGGATGCAGAGGCTGAGGGTCACTGTATCTGAAGCCCTTCGAGATAAAG	1091
QY	1906	CGCGTTGACCGGCC 1918	
Db	1092	CGCGTTGACCGGCC 1104	

XX	RESULT 11
XX	AAC76397
ID	AAC76397 standard; cDNA; 598 BP.
XX	
AC	AAC76397;
XX	
DT	08-FEB-2001 (first entry)
DE	Human ORFX ORF1952 polynucleotide sequence SEQ ID NO:3903.
XX	
KW	Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW	vulnery; antisociatic; antiparkinsonian; nootropic; neuroprotective;
KW	anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW	immunostimulant; thrombolytic; coagulant; vasotropic; antiabiotic;
KW	hypotensive; dematological; immunosuppressive; antiinflammatory;
KW	antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW	antianeatic; gene therapy; cancer; proliferative disorder; hypertension;
KW	neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW	cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW	cholesterol ester storage; systemic lupus erythematosus; infection;
KW	severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW	bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW	thrombosis; contraceptive; ss.
XX	
OS	Homo sapiens.
XX	
PN	MO200058473-A2.
XX	
XX	05-OCT-2000.
XX	
PF	31-MAR-2000; 2000WO-US08621.
XX	
PR	31-MAR-1999; 99US-0127607.
PR	02-APR-1999; 99US-0127636.
PR	05-APR-1999; 99US-0127728.
PR	30-MAR-2000; 2000US-0540763.
XX	
XX	(CURA-) CURAGEN CORP.
XX	

PI Shimkets RA, Leach M;  
XX  
XX  
DR WPI, 2000-602362/57,  
DR P-PSDB; AAB42188.  
XX  
XX Novel nucleic acids and peptides derived from open reading frame X  
PT useful for treating e.g. cancers, proliferative disorders,  
XX neurodegenerative disorders and cardiovascular disease -  
XX  
PS Claim 5, Page 3050, 5507pp, English.

CC AACT4446 to AACT7606 encode the proteins given in AAB40237 to AAB53397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cycostatic; hepatoprotective; vulnery;  
CC antiprolastic; antiparkinsonian; nootropic; neuroprotective;  
CC osteoplastic; anticonvulsant; antirhectic; immunosuppressant;  
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
CC antihlammatory; antibacterial; antiviral; antifungal; antineumatic;  
CC antihydroid; and antianemic. The sequences can be used for determining  
CC the presence of or predisposition to, or preventing or treating  
CC pathological conditions associated with an ORFX-associated disorder. The  
CC nucleic acids can be used to express ORFX proteins in gene therapy  
CC vectors. The proteins and nucleic acids may be used to treat cancers,  
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
CC nocturnal haemoglobinuria, antihlammatory disease), to enhance  
CC coagulation; to inhibit thrombosis; and as a contraceptive.

Sequence: 598 BP; 84 A; 189 C; 192 G; 133 T; 0 other;

Query Match 20.4%; Score 396; DB 21; Length 598;

Matches 396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	296	GGCGGCGGGAGAGCGCGGTGTGTGCCTGCGCCAGACAGAGAGCGGGAGCGGCT	355
Db	1	GGCGGCGGGAGAGCGCGGTGTGTGTGCCTGCGCCAGAGAGCGGGAGCGGCT	60
Qy	356	GCGTTGACCTCCGCGAGAGAGTGGGAAACAATGAGTCATCTTATGCGCTTGAAGCTTG	415
Db	61	GCGTTGACCTCCGCGAGAGAGTGGGAAACAATGAGTCATCTTATGCGCTTGAAGCTTG	120
Qy	416	GCGAGTGTGGCTCGGTGCGGGCCTTTGCACTGCTTCTGAGCTCTGAGCCACGGTTG	475
Db	121	GCGAGTGTGGCTCGGTGCGGGCCTTTGCACTGCTTCTGAGCTCTGAGCCACGGTTG	180
Qy	476	GACATCCGCATCCACAATGACGGATATCAAGTCCCTGTGCGCGGAGCCCGTAGAGCGTTTAA	535
Db	181	GACATCCGCATCCACAATGACGGATATCAAGTCCCTGTGCGCGGAGCCCGTAGAGCGTTTAA	240
Qy	536	CTGCTGCTTCGGGTGAACCATATCGATCCCTTCTGTGACACATCTGCTGCTGCTTGC	595
Db	241	CTGCTGCTTCGGGTGAACCATATCGATCCCTTCTGTGACACATCTGCTGCTTGC	300
Qy	596	CTGAAGCATGTGCCCTTAGCGCGCGGTGTGTGAGCTTCAGCTGCCCATGTGCGGGA	655
Db	301	CTGAAGCATGTGCCCTTAGCGCGCGGTGTGTGAGCTTCAGCTGCCCATGTGCGGGA	360
Qy	656	CGCTTGACCTCAAAACGCTGAGACCGGCCAAGGTG	691
Db	361	CGCTTGACCTCAAAACGCTGAGACCGGCCAAGGTG	396

RESULT 12
AA66494
ID AA66494 standard, cDNA, 397 bp
XX
AC AA66494;



XX 09-APR-2001 (first entry)  
XX Novel human polynucleotide, SEQ ID NO: 2250.  
XX  
XX Human; cytostatic; gene therapy; colon cancer; prostate cancer;  
XX breast cancer; lung cancer; cancer detection; 85.  
XX Homo sapiens.  
XX WO200102568-A2.  
XX  
XX 11-JAN-2001.  
XX  
XX 30-JUN-2000; 2000WO-US18374.  
XX  
XX 02-JUL-1999; 99US-0142310.  
XX 02-JUL-1999; 99US-0142311.  
XX  
XX (CHIR) CHIRON CORP.  
XX (HYSB-) HYSBQ INC.  
XX Williams LT, Sacobedo J, Innis MA, Garcia PD, Klinger J, Kasam A,  
XX Reinhard C, Randazzo F, Kennedy GC, Pot D, Iamson G, Drmanac R,  
XX Cikenjakov R, Drmanac S, Dickson M, Labat I, Leishowitz D,  
XX Kita D, Garcia V, Jones LW, Strache-Crain B,  
XX WPI; 2001-091805/10.  
XX  
XX Library of polynucleotides for diagnosing a cancerous state of a  
XX mammalian cell and detecting cancer, particularly of the colon or  
XX prostate, comprises 3351 human polynucleotide sequences -  
XX  
XX Claim 9; Page 873; 1046pp; English.  
XX  
XX The present sequence is one of 3351 sequences in a library of human  
XX polynucleotides. The library is used to detect differentially expressed  
XX genes correlated with a cancerous state of a mammalian cell and can  
XX detect colon, prostate, breast and lung cancer. The library can be used  
XX to produce probes for detection of mRNA and to produce additional copies  
XX of the polynucleotides. The probes can be used for chromosome mapping of  
XX the polynucleotide and for detection of transcription levels. Ribozymes  
XX or antisense oligonucleotides can be generated. The polynucleotides and  
XX their gene products are used as genetic or biochemical markers (e.g. in  
XX blood or tissues) that will detect the earliest changes along the  
XX carcinogenesis pathway and/or monitor the efficacy of therapies and  
XX preventive interventions. The polynucleotides, polypeptides and  
XX antibodies against them can be used in pharmaceutical compositions to  
XX treat the cancers and proliferative disorders such as neoplasia,  
XX dysplasia and hyperplasia.  
XX  
XX Sequence 397 BP; 74 A; 116 C; 125 G; 82 T; 0 other;  
SQ  
Query Match 10.3%; Score 200; DB 22; Length 397;  
Best Local Similarity 100.0%; Pred. No. 3.7e-68;  
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 716 TATGCTGACACTAAGTGGCTAATGTAATGTTTCCCGAGAGCTGGCAACAGCTGAG 775  
DB 31 TATGCTGACACTAAGTGGCTAATGTAATGTTTCCCGAGAGCTGGCAACAGCTGAG 90  
QY 776 GCCACTGCGCTGACCTGCTATGACGCCACCCAGGCGCTGTGAACCTGGAGCTGTTCTG 835  
DB 91 GCCACTGCGCTGACCTGCTATGACGCCACCCAGGCGCTGTGAACCTGGAGCTGTTCTG 150  
QY 836 CGCATATGCTCTGATGAGTGGCGGCCCACTTTGGCGCCCATGAGTGGCTGGCTCGG 895  
DB 151 CGCATATGCTCTGATGAGTGGCGGCCCACTTTGGCGCCCATGAGTGGCTGGCTCGG 210  
QY 896 GCACCAAGAGGGGCTGCCCA 915  
DB 211 GCACCAAGAGGGGCTGCCCA 230

RESULT 13  
AAT19666  
ID AAT19666 standard; cDNA to mRNA; 186 BP.  
XX  
XX AAT19666;  
XX  
XX 28-JUN-1996 (first entry)  
XX  
XX Human gene signature HUMGS00736.  
XX  
XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;  
XX human; cloning; mapping; non-biased library; diagnosis; detection;  
XX cell typing; abnormal cell function; 88.  
XX Homo sapiens.  
XX  
XX WO9514772-A1.  
XX  
XX 01-JUN-1995.  
XX  
XX 11-NOV-1994; 94WO-JP01916.  
XX  
XX 12-NOV-1993; 93JP-0355504.  
XX  
XX (MATS/) MATSUBARA K.  
XX (OKUB/) OKUBO K.  
XX  
XX Matsubara K, Okubo K;  
XX  
XX WPI; 1995-206931/27.  
XX  
XX Identifying gene signatures in 3'-directed human cDNA library - e.g.  
XX for diagnosis of abnormal cell function, by preparing cDNA that  
XX reflects relative abundance of corresp. mRNA in specific human  
XX tissues  
XX  
XX Claim 1; Page 445; 2245pp; Japanese.  
XX  
XX A single-stranded DNA (or its complementary strand or the corresp.  
XX double-stranded DNA) which comprises one of the 7837 "GS" sequences  
XX given in AAT19001-726837 and which is able to hybridise to part of  
XX human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)  
XX sequences were obtained from 3'-directed cDNA libraries prepared  
XX from various human tissues; synthesis of cDNA was initiated from the  
XX 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
XX untranslated sequence is unique to a particular mRNA species, almost  
XX all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library  
XX is constructed so as to reflect accurately the relative abundance of  
XX different mRNAs in the particular tissue from which it was derived.  
XX The appearance frequency of a given GS in a cDNA library can be  
XX determined (esp. using primers and probes derived from the GS  
XX sequences) as a means of diagnosing abnormal cell function or for  
XX recognising different cell types.  
XX  
XX Sequence 186 BP; 42 A; 42 C; 67 G; 31 T; 4 other;  
SQ  
Query Match 4.7%; Score 92; DB 16; Length 186;  
Best Local Similarity 100.0%; Pred. No. 4.1e-26;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1744 AACCCAAAGGCTCTGAGGCCGAGGCGGACCTGCCCTAAGATGGTGTGAGAGTGTGAG 1803  
DB 8 AACCCAAAGGCTCTGAGGCCGAGGCGGACCTGCCCTAAGATGGTGTGAGAGTGTGAG 67  
QY 1804 GGCAGGGCAGCTGTGATCCAGAGTGGCCCATGG 1835  
DB 68 GGCAGGGCAGCTGTGATCCAGAGTGGCCCATGG 99  
RESULT 14  
ABK64203/c  
ID ABK64203 standard; DNA; 175 BP.



